

Universität Hildesheim  
Fachbereich III – Informations- und Kommunikationswissenschaften  
Institut für Informationswissenschaft und Sprachtechnologie



# ADAPTIERBARE ONLINE-NUTZERBEFRAGUNGEN ZUR BEURTEILUNG VON WEBSITES

Magisterarbeit zur Erlangung des akademischen Grades Magistra Artium des Internationalen  
Informationsmanagements

Franziska Leithold  
Hildesheim, April 2010  
(Matr.Nr. 196155)

Erstgutachterin: Prof. Dr. Christa Womser-Hacker  
Zweitgutachter: Dr. Thomas Mandl

*[As you set out for Ithaka,  
hope your road is a long one,  
full of adventure, full of discovery.]*  
*(Konstantinos Kavafis)*

## Danksagung

An erster Stelle möchte ich mich bei meinen Betreuern für die Unterstützung und Hilfestellung bedanken, hierbei besonders Frau Prof. Dr. Womser-Hacker für die Korrekturen und inhaltlichen Hinweise und Herr Dr. Mandl für die Bereitschaft, Teile der Arbeit mit Seminarteilnehmern zu testen.

Mein besonderer Dank gilt Ben Heuwing, der für Fragen immer ein offenes Ohr hatte, und mich von Anfang bis Ende mit Anregungen und Hilfestellung unterstützt hat.

Da man eine solche Arbeit nie ohne Familie und gute Freunde schaffen kann, danke ich darüber hinaus Thorsten für seine Geduld und Leila für ihre guten Ratschläge, und allen Freunden, die mich unterstützt haben, sowie meiner Mutter und meiner Tante für die Korrekturen.

Ich möchte diese Arbeit meinem Vater widmen, der uns allen sehr fehlt.



## **Zusammenfassung**

Gegenstand der vorliegenden Magisterarbeit ist die Entwicklung adaptierbarer Fragebögen für Benutzerbefragungen zur Erhebung der Gebrauchstauglichkeit von Websites unterschiedlicher Kategorien und einer entsprechenden Anwendung für diesen Zweck. Hierzu wird ein Klassifikationsschema entworfen, welches insgesamt sechs Anwendungskontexte des World Wide Web abdeckt: E-Commerce, Entertainment, Social Software, Information, Repräsentation und E-Learning. Für diese Kategorien werden aus der Literatur entsprechende Usability-Kriterien ermittelt, und aus ihnen im Anschluss Items für kategoriespezifische Fragebögen zur Erhebung der Gebrauchstauglichkeit generiert. Die Items werden vor dem Einsatz unter Endnutzern von Experten und Studierenden getestet, um zu ermitteln, ob sie in Qualität und Vollständigkeit für eine Nutzerbefragung geeignet sind. Anschließend erfolgen der Test der kompletten Itemsets jeder Kategorie in der konkreten Usability-Evaluation und die Auswertung der Antwortergebnisse. Die Arbeit schließt mit der prototypischen Erstellung der Anwendung ab, die die Grundlage für eine spätere Implementierung bilden wird.

### **Schlüsselbegriffe:**

Usability, Web-Usability, Evaluation, Fragebogen, Nutzerbefragung, Gebrauchstauglichkeit, Adaptierbarkeit, Website-Klassifikation, Qualität, Qualitätssicherung, User-Centered Design

## Abstract

This M. A. Thesis deals with the development of adaptive questionnaires for the user-centered evaluation of website usability and an appropriate web application for this purpose. In this matter, a classification scheme for six various website contexts is generated: E-Commerce, Entertainment, Social Software, Information, Homepages (organizational/personal) and E-Learning. For each category, specific usability criteria based on recent literature are identified. These will be the groundwork for the development of category-based 5-level Likert scale items of which the later questionnaires will be composed. The questions will be tested and evaluated by usability-experts and students to ensure their quality and usefulness. After that, a final test with end-users is done using the complete question set of each category to evaluate a corresponding website followed by the analysis of the results. In a final step, a prototype for the web application is developed which will build the basis for the upcoming implementation.

Keywords: usability, web-usability, evaluation, questionnaires, surveys, adaptivity, web-classification, quality, quality assurance, user-centered design

## INHALT

A	Tabellenverzeichnis.....	X
B	Abbildungsverzeichnis .....	XII
1	Einleitung .....	1
1.1	Das Projekt <a href="http://www.usability-toolkit.de">www.usability-toolkit.de</a> .....	1
1.2	Motivation und Aufbau der Arbeit.....	1
2	Usability.....	2
2.1	Web-Usability.....	2
2.2	User Experience .....	4
2.3	Dimensionen von Web-Usability.....	6
2.4	Usability-Probleme.....	7
2.5	Fazit.....	8
3	Methoden der Usability-Evaluation .....	9
3.1	Expertenorientierte Methoden der Usability-Evaluation .....	9
3.2	Benutzerorientierte Methoden der Usability-Evaluation .....	10
3.3	Fazit.....	10
4	Empirische Usability-Evaluation mit Fragebögen.....	10
4.1	Fragebogaufbau .....	11
4.2	Computervermittelte Befragungen.....	12
4.3	Fragebogenkonstruktion und Gestaltrichtlinien zur Online-Erstellung .....	13
4.4	Teilnehmerkreis & Stichprobenwahl.....	14
4.5	Fazit.....	14
5	Testtheorie & Fragebogenverfahren.....	15
5.1	Fragebogenkonstruktion .....	16
5.2	Haupt- und Nebengütekriterien.....	17
5.2.1	Hauptgütekriterien.....	17
5.2.2	Nebengütekriterien .....	18
5.3	Evaluierung der Methodik des Online – Fragebogenverfahrens .....	19
5.3.1	Praktische Relevanz.....	20

5.3.2	Wissenschaftliche Relevanz .....	22
5.4	Fazit.....	23
6	Bestehende Fragebögen zur Online-Nutzerbefragung .....	24
6.1	QUIS (Questionnaire for User Interface Satisfaction) .....	24
6.2	SUMI (Software Usability Measurement Inventory).....	25
6.3	Isometrics.....	26
6.4	IsoNorm 9241.....	26
6.5	WAMMI (Website Analysis and Measurement Inventory) .....	27
6.6	AttrakDiff.....	27
6.7	UEQ (User Experience Questionnaire) .....	28
6.8	Zusammenfassung .....	29
6.9	Fazit.....	29
7	Klassifikation von Websites und unterschiedliche Anwendungskontexte .....	30
7.1	Klassifikation unterschiedlicher Website-Typen .....	31
7.1.1	Methodisches Vorgehen .....	31
7.1.2	Bestehende Klassifikationsansätze.....	31
7.1.3	Zusammenfassung.....	34
7.2	Entwicklung einer eigenen Klassifikation zum Einsatz im Fragebogengenerator .....	35
7.3	Zusammenfassung .....	37
7.4	Spezifische Usability-Anforderungen in den unterschiedlichen Kategorien .....	38
7.4.1	Methodisches Vorgehen .....	39
7.4.2	E-Commerce .....	39
7.4.3	Entertainment .....	43
7.4.4	Information .....	44
7.4.5	Portale, Communities, Social Software .....	46
7.4.6	Präsentation .....	49
7.4.7	E-Learning.....	50
7.5	Zusammenfassung .....	52
7.6	Einordnung der ermittelten Kriterien .....	52
7.7	Fazit.....	55

8	Erstellung der Fragebogenitems .....	55
8.1	Methodisches Vorgehen zu Skalenwahl und Itemtyp.....	55
8.2	Revision des Itempools durch Experten.....	58
8.3	Revision des Itempools durch Studierende der Informationswissenschaft.....	59
8.4	Relevanzerrechnung .....	60
8.5	Freitextkommentare der Experten .....	62
8.6	Freitextkommentare der Studenten .....	64
8.7	Überarbeitung des Itempools .....	65
8.8	Pretest der Items mit Endnutzern.....	67
8.9	Fazit.....	68
9	Benutzertest der Fragebögen .....	68
9.1	Erstellung der Fragebögen in Limesurvey.....	68
9.2	Online-Pretest und Durchführung des endgültigen Benutzertests.....	69
9.3	Server- und DNS-Probleme mit LimeSurvey .....	69
9.4	Durchführung der Umfrage.....	70
9.5	Rücklaufquote .....	72
9.6	Fazit.....	73
10	Statistische Auswertung und Güte des Fragebogens.....	73
10.1	Deskriptive Statistik über die einzelnen Items.....	73
10.1.1	E-Commerce: <a href="http://www.vv-computer.de">www.vv-computer.de</a> .....	74
10.1.2	Entertainment: <a href="http://www.youtube.com">www.youtube.com</a> .....	75
10.1.3	E-Learning: <a href="http://www.nocomprendo.es/spanischkurs_online">www.nocomprendo.es/spanischkurs_online</a> .....	77
10.1.4	Information: <a href="http://www.spiegel-online.de">www.spiegel-online.de</a> .....	78
10.1.5	Social Software: <a href="http://www.facebook.com">www.facebook.com</a> .....	79
10.1.6	Repräsentation: <a href="http://www.bhpeg.de">www.bhpeg.de</a> .....	81
10.2	Zusammenfassung .....	81
10.3	Diskussion der Fragebögen unter den allgemeinen Gütekriterien .....	82
10.3.1	Reliabilität .....	82
10.3.2	Validität.....	84
10.3.3	Nebengütekriterien.....	85



10.4	Fazit.....	85
11	Prototypische Erstellung des Fragebogens .....	86
11.1	Anforderungsanalyse .....	86
11.2	Architektur .....	87
11.3	Programmfunktionen.....	88
11.4	Aktivitätsdiagramm .....	92
11.5	Gestaltungsgrundsätze nach ISO-9241 für den Prototyp .....	95
11.6	Erstellung der Benutzeroberfläche .....	96
11.7	Seitenstruktur des Prototypen.....	98
11.7.1	Startseite.....	99
11.7.2	Begrüßung.....	100
11.7.3	Neuen Bogen erstellen .....	101
11.7.4	Items der Hauptkategorie .....	102
11.7.5	Zusatzkategorien.....	103
11.7.6	Vorschau .....	104
11.7.7	Daten und Einstellungen.....	105
11.7.8	Auswertung.....	106
11.8	Zusammenfassung .....	107
12	Fazit und Ausblick .....	107
13	Literaturverzeichnis .....	109
	Appendix A – Gegenüberstellung der Kategorisierungen von Thielsch/Hong & Kim .....	117
	Appendix B - Zuordnung der ermittelten Usability-Kriterien zu den einzelnen Gestaltungsgrundsätzen .....	118
	Appendix C – Itemsätze.....	124
	Appendix D – Anschreiben an die Experten und Studierenden .....	132
	Appendix E: Experten und Studierenden Relevanzbewertung .....	134
	Appendix F: Auflistung der überarbeiteten Items mit Relevanzbewertung .....	134
	Appendix G – Histogramme und Tabellen der SPSS-Auswertung.....	143
	Appendix H – Prototyp .....	143
	Appendix I – Halbierungsreliabilität der Kategorie Entertainment.....	144
	Eigenständigkeitserklärung nach §26 Abs. 6 der Magisterprüfungsordnung .....	145

## A TABELLENVERZEICHNIS

Tabelle 1: Dimensionen von Web-Usability .....	7
Tabelle 2: Praktische Relevanz von Fragebogenverfahren in der Usability-Evaluation .....	22
Tabelle 3: Wissenschaftliche Relevanz von Fragebogenverfahren in der Usability-Evaluation .....	23
Tabelle 4: Klassifikationsansätze nach unterschiedlichen Autoren .....	34
Tabelle 5: Klassifikationen von Websites nach Thielsch (2008) /Hong&Kim (2004) .....	35
Tabelle 6: Beispielwebsites für die ermittelten Kategorien .....	37
Tabelle 7: Usability-Kriterien der Kategorie E-Commerce .....	42
Tabelle 8: Usability-Kriterien der Kategorie Entertainment.....	44
Tabelle 9: Usability-Kriterien der Kategorie Information.....	46
Tabelle 10: Usability-Kriterien der Kategorie Social Software .....	48
Tabelle 11: Usability-Kriterien der Kategorie Repräsentation .....	50
Tabelle 12: Usability-Kriterien der Kategorie E-Learning .....	52
Tabelle 13: Zuordnung beispielhafter Usability-Kriterien zu den Grundsätzen der Dialoggestaltung.....	54
Tabelle 14: Zuordnung beispielhafter Usability-Kriterien zu den Grundsätzen der Gestaltung von Multimedia- Anwendungen .....	55
Tabelle 15:Anzahl an Items in den jeweiligen Kategorien .....	57
Tabelle 16: Beispielhafte Itemformulierungen .....	57
Tabelle 17: Anzahl Expertenurteile für die jeweiligen Kategorien .....	59
Tabelle 18: Anzahl Studierendenurteile für die jeweiligen Kategorien.....	60
Tabelle 19: Mittelwerte für die Relevanzscores.....	61
Tabelle 20: Mittelwerte der endgültigen Relevanzscores der einzelnen Kategorien .....	62
Tabelle 21: Auszug der Expertenkommentare für den Bereich E-Commerce und E-Learning .....	63
Tabelle 22: Auszug der Studierendenkommentare für den Bereich E-Commerce und Entertainment.....	64
Tabelle 23: Anzahl der Items vor und nach der Bewertung durch Experten und Studierende.....	65
Tabelle 24: Beispielhafter Vergleich von ursprünglichen und überarbeiteten Items .....	66
Tabelle 25: Rücklaufquote der Umfrage nach Ablauf des Zeitraums von 19 Tagen .....	72
Tabelle 26: E-Commerce - Umgepolte Items .....	75
Tabelle 27: E-Commerce - Items mit auffälliger Antwortverteilung .....	75

Tabelle 28: Entertainment - Umgepolte Items .....	76
Tabelle 29: Entertainment – Items mit hoher Nonresponse-Quote .....	76
Tabelle 30: Entertainment – Items mit auffälliger Antwortverteilung.....	76
Tabelle 31: E-Learning – Umgepolte Items .....	78
Tabelle 32: E-Learning – Auffällige Antwortverteilungen .....	78
Tabelle 33: Information - Umgepolte Items.....	79
Tabelle 34: Information - Auffällige Antwortverteilung.....	79
Tabelle 35: Social Software – Umgepolte Items .....	80
Tabelle 36: Social Software – Auffällige Antwortverteilungen .....	80
Tabelle 37: Repräsentation – Umgepolte Items .....	81
Tabelle 38: Repräsentation – Auffällige Antwortverteilungen .....	81
Tabelle 39: Korrelationskoeffizienten für die sechs Kategorien .....	84
Tabelle 40: Legende des Aktivitätsdiagramms.....	93
Tabelle 42: Verlinkung der HTML-Seite auf den Buttons.....	99

## B ABBILDUNGSVERZEICHNIS

Abbildung 1: QUIS-Skala.....	25
Abbildung 2: SUMI – Skala, UK-Sample.....	25
Abbildung 3: IsoMetrics-Skala.....	26
Abbildung 4: IsoNorm 9241 .....	27
Abbildung 5: WAMMI – Skala .....	27
Abbildung 6: Semantisches Differential des AttrakDiff.....	28
Abbildung 7: User Experience Questionnaire .....	28
Abbildung 8: Struktur und Aufbau von Websites nach Thielsch 2008 .....	30
Abbildung 9: Kategorisierung des Web nach Hong, Kim (2004) .....	32
Abbildung 10: Kategorisierung des Web nach Levialdi, DeMarsico (2004).....	33
Abbildung 11: Ausschnitt aus einem Experten-Testbogen .....	58
Abbildung 12: Ausschnitt aus einem Pretest-Benutzerbogen .....	67
Abbildung 13: Startseite der Umfrage mit Anschreiben und Datenschutzerklärung.....	70
Abbildung 14: Beispielseite der erstellten Umfrage .....	71
Abbildung 15: Letzte Seite der Umfrage mit Verabschiedung und Kommentarfeld.....	72
Abbildung 16: Client-Server Architektur des Prototypen.....	87
Abbildung 17: Aktivitätsdiagramm des Fragebogengenerators.....	94
Abbildung 18: Benutzeroberfläche des Programms AXURE .....	97
Abbildung 19: E-Quest – Startseite .....	99
Abbildung 20: E-Quest – Begrüßungsbildschirm nach dem Login .....	100
Abbildung 21: E-Quest – Erstellung eines neuen Bogens und Kategoriewahl .....	101
Abbildung 22: E-Quest – Items der Hauptkategorie .....	102
Abbildung 23: E-Quest – Items der Zusatzkategorie.....	103
Abbildung 24: E-Quest – Vorschau des Social Software-Bogens .....	104
Abbildung 25: E-Quest – Daten und Einstellungen .....	105
Abbildung 26: E-Quest – Auswertung der Kategorie E-Commerce.....	106

## 1 EINLEITUNG

### 1.1 DAS PROJEKT WWW.USABILITY-TOOLKIT.DE

Die vorliegende Arbeit ist im Rahmen des Projektes „Toolkit für B-U-I: Barrierefreiheit, Usability, Internationalisierung“ (im Folgenden abgekürzt als BUI-Projekt bezeichnet) an der Universität Hildesheim entstanden. Das BUI-Projekt bildet in Zusammenarbeit mit dem Unternehmen usability.de in Hannover eine wichtige Initiative mit der Zielsetzung, kleine und mittelständische Unternehmen, die über kein oder lediglich marginales Wissen im Fachgebiet Usability verfügen, mit den Grundlagen von Usability, User Centered Design und User Experience sowie ihrer Evaluation vertraut zu machen und die dafür notwendigen Werkzeuge auf der Plattform [www.usability-toolkit.de](http://www.usability-toolkit.de) online zur Verfügung zu stellen. Ein wichtiger Aspekt der Plattform ist die explizite Nichtvoraussetzung von Expertenwissen bei ihren Nutzern, sowie deren Möglichkeit, sich praxisorientiert und umfassend über methodische Vorgehensweisen bei der Analyse, Konzeption und Umsetzung von Webprojekten zu informieren sowie spezielle Tools zu deren Test und Evaluation in Anspruch zu nehmen.

Benutzerbefragungen stellen einen wesentlichen Teilbereich der Evaluation von Software im Allgemeinen und Webprojekten im Speziellen dar. Sie sind in der Lage, ein fundiertes und umfassendes Bild von der Gebrauchstauglichkeit und – bei entsprechender Ausrichtung – des qualitativen Nutzungserlebnisses (User Experience) aus der Benutzerperspektive zu liefern und ergänzen so heuristische Expertenmethoden zur Usability-Evaluation in wichtiger Weise.

### 1.2 MOTIVATION UND AUFBAU DER ARBEIT

Die vorliegende Arbeit beschäftigt sich mit der Erstellung eines Tools zur kontextspezifischen Generierung von Online-Nutzerbefragungen zur Beurteilung von Websites. Den Nutzern des Tools soll die Möglichkeit geboten werden, ohne psychologisches Expertenwissen und Vorkenntnisse im Bereich der Fragebogenkonstruktion die Erstellung eines auf ihre jeweilige Seite angepassten Fragebogens zur Erhebung von Daten zur Usability ihres Webauftrittes oder -projektes vorzunehmen. Wichtig ist hierbei, dass keine allgemeinen (s. Anmerkung zur Terminologie) Usability-Probleme mit dem Fragebogen ermittelt werden sollen, da Erhebungen dieser Art bereits existieren (s. Abschnitt 6) und dies auch durch eine Expertenevaluation möglich wäre. Vielmehr sollen konkrete Schwierigkeiten und Hindernisse des jeweiligen Anwendungskontextes ermittelt werden. Das Fehlen allgemeiner Usability-Kriterien in der vorliegenden Arbeit ist entsprechend keine Auslassung, sondern bewusst gewollt.

Hierfür soll im theoretischen Teil der Arbeit (Kapitel 2 – 6) nach einer allgemeinen Einführung in die Thematik Usability und User Experience (Kapitel 2), ihrer Evaluation (Kapitel 3) sowie der allgemeinen Testtheorie (Kapitel 4) auf die Einordnung des Bereichs Benutzerbefragungen in der Usability-Evaluation (Kapitel 5) eingegangen werden, wobei bereits bestehende Fragebögen zur Evaluation von Software und Webprojekten beschrieben und analysiert werden (Kapitel 6). Im

Anschluss (Kapitel 7) wird in Anlehnung an bereits bestehende Klassifikationsansätze eine eigene Kategorisierung unterschiedlicher Anwendungskontexte von Websites vorgenommen. Im praktischen Teil der Arbeit (Kapitel 8 – 9) sollen für die jeweiligen Websitekategorien Itempools (Gesamtheit der Items für die jeweilige Kategorie) erstellt werden, aus denen später die Selektion von Items zur Generierung des kontextspezifischen Fragebogens erfolgen kann (Kapitel 8). Diese Items sollen mit Usability-Experten und Studierenden der Informationswissenschaft getestet werden, um ihre Validität zu gewährleisten, sowie im Anschluss überarbeitet werden. Die so entstandenen fertigen Itempools sollen schließlich mit Endnutzern als Online-Fragebögen durchgeführt sowie im Anschluss statistisch ausgewertet und auf ihre Güte geprüft werden (Kapitel 9).

Der abschließende Teil der Arbeit (Kapitel 10) beschäftigt sich mit einer prototypischen Erstellung des Tools zur Fragebogengenerierung. Hierbei soll eine Anforderungsanalyse zusammen mit einer begrenzt interaktiven Oberfläche (sog. „Clickdummy“) erstellt werden. Diese bilden zusammen die Grundlage für eine spätere Implementierung. Kapitel 11 wird sich mit einer kurzen Zusammenfassung und dem Ausblick auf weiteren Arbeitsbedarf beschäftigen.

Anmerkung der Autorin zur Terminologie:

Insofern nicht explizit anders dargelegt, werden die Bezeichnungen „Anwendung“, „Tool“ und „Fragebogengenerator“ im Zusammenhang mit dem zu erstellenden Prototypen synonym verwendet. Die Begriffe „Nutzer“, „User“ und „Anwender“ werden im Folgenden sowohl im Zusammenhang mit Software als auch mit Websites synonym gebraucht. Insofern von „allgemeiner Usability“ oder „allgemeinen Usability-Kriterien“ gesprochen wird, kann diese als Gebrauchs-tauglichkeit nach den allgemeinen Grundsätzen der Dialoggestaltung (DIN EN ISO 9241), wie sie in bestehenden Bögen zur Usability-Evaluation beurteilt wird, verstanden werden und ist von den in der folgenden Arbeit erstellten, kategoriespezifischen Usability-Kriterien zu trennen.

## 2 USABILITY

### 2.1 WEB-USABILITY

Mit der Entstehung und dem rapiden Wachstum des World Wide Web (WWW) ist sowohl die Anzahl der Websites als auch die der Internetnutzer seit den 90er Jahren in einem stetigen Anstieg begriffen. Ursprünglich als reines Medium für den vereinfachten Austausch von Daten und Information am Europäischen Labor für Teilchenphysik entworfen [vgl. Thielsch 2008:3], hat sich das Internet heutzutage zu einem globalen Kommunikations- und Informationsmedium gewandelt, zu dem ein stetig wachsender Teil der Menschheit Zugriff hat. Zwischen 43,2 [s. AGOF 2009:5] und 46,3 [s. TNS Infratest & Initiative D21 2009:10] Milliarden Deutsche ab 14 Jahren nutzen das Internet (ca. 63,1% der Bevölkerung). Von diesen nutzen wiederum ca. 71,3% [s. TNS

Infratest & Initiative D21 2009:58] das Web über einen Breitbandanschluss. Die Anzahl indexierter Websites (Stand: 31. März 2010) beträgt nach Angaben der Betreiber von [www.worldwidewebsite.com](http://www.worldwidewebsite.com) mindestens 20.48 Milliarden, die Anzahl der deutschen registrierten Domains beträgt derzeit ca. 13,5 Millionen<sup>1</sup>.

Die Nutzungsmöglichkeiten, Zielgruppen und Anwendungskontexte des Internets sind vielfältiger Natur. Laut einer Studie der Arbeitsgemeinschaft Online-Forschung e.V. werden E-Mail-Versand (88.6%), Informationsrecherche (86.8%), Nachrichten (65.1%) sowie Online-Shopping (61.9%) als häufigste Nutzungsarten genannt. Auch Chats und Foren (41.7%) sowie Messenger (40.1%) sind wichtige Kommunikationsmittel, die über das World Wide Web genutzt und betrieben werden [s. AGOF 2009:10].

Vor diesem Hintergrund ist der Faktor Usability im Web von entscheidender Bedeutung geworden. Der Erfolg oder Misserfolg von Webauftritten hängt im Wesentlichen von ihrer Gebrauchs-tauglichkeit und dem subjektiven Nutzungserlebnis des Users ab. Usability wird nach DIN EN ISO 9241 (im Folgenden kurz als ISO 9241 bezeichnet) als

*„...das Ausmaß, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele effektiv, effizient und mit Zufriedenheit zu erreichen.“*

definiert [von Gizycki, Beier 2002:2].

Effektivität wird hierbei beschrieben als

*„...die Genauigkeit und Vollständigkeit, mit der Benutzer ein bestimmtes Ziel erreichen.“*

[ISO 9241-11 96:8 in von Gizycki, Beier 2002:2]

Effizienz wird bei den Autoren in Anlehnung an ISO-9241 dadurch definiert, dass

*„ ...der Nutzer zum Erreichen seines Ziels den geringstmöglichen Einsatz benötigt[...]. Der Aufwand muss für den Nutzer geringer als der durch die Website erreichte Ertrag sein oder ihm entsprechen, sonst würde der Nutzer den Vorgang abbrechen.“*

[a.a.O.:3]

Zufriedenheit als Kenngröße der Usability ist eine schwieriger zu quantifizierende Eigenschaft, da

*„...sie von zahlreichen verschiedenen Erwartungen des Nutzers abhängt. ‘Maße der Zufriedenheit beschreiben die Beeinträchtigungsfreiheit und die Akzeptanz der Nutzung. [...] Maße der*

---

<sup>1</sup> Stand: 31. März 2010, <http://www.denic.de/hintergrund/statistiken.html>

*Zufriedenheit können sich auf Einstellungen beziehen, ein Produkt zu benutzen, oder auf das Benutzerurteil über Aspekte wie Effizienz, Nützlichkeit oder Lernförderlichkeit.“*

[ISO 9241-11 96:8 in von Gizycki, Beier 2002:3]

Zum anderen muss die Kontextabhängigkeit des Konzeptes Usability berücksichtigt werden. Die Autoren halten fest, dass nach ISO-9241

*„ ...Effizienz, Effektivität und Zufriedenheit nicht durch eindeutige und dauerhafte Maße festgehalten werden können: ‘Weil die relative Bedeutung dieser Komponenten der Gebrauchstauglichkeit sowohl vom Nutzungskontext abhängt als auch vom Zweck, für den die Gebrauchstauglichkeit zu beschreiben ist, gibt es keine allgemeine Regel dafür, wie Maße ausgewählt und kombiniert werden sollen.’ Das bedeutet, dass Usability immer zweckbezogen betrachtet werden muss.“*

[ISO 9241-11 96:8 in von Gizycki, Beier 2002:3]

Festzuhalten ist, dass die drei Kenngrößen der Usability Effektivität, Effizienz und Zufriedenheit stark interdependent sind. Vor allem der subjektive Faktor Zufriedenheit ist in direkt kausaler Beziehung davon abhängig, ob ein System in der Lage ist, die Erwartungen seiner Benutzer an seine Leistungsfähigkeit zu erfüllen [vgl. a.a.O.:4]. Thissen und Schweibenz verweisen bezüglich der Gebrauchstauglichkeit im Internet auf eine wesentlich geringere Toleranzschwelle für Usability-Probleme bei Nutzern. Dies hat zur Konsequenz, dass Nutzungsschwierigkeiten und -probleme bei Webauftreten mit hoher Wahrscheinlichkeit dazu führen, dass der Nutzer die entsprechende Seite mit dem Ziel verlässt, eine alternative und unter Umständen besser gestaltete Website aufzusuchen, um sein jeweiliges Vorhaben umzusetzen.

Die Autoren weisen im Zusammenhang mit der Vielfalt der Nutzungsmöglichkeiten des Internets ebenfalls auf die Schwierigkeit eines festgelegten Konzeptes für Web Usability hin und empfehlen entsprechend eine ex-Ante-Definition zum einen der Nutzungsart und zum anderen der Zielgruppe, da unterschiedliche Anwendungskontexte unter Umständen unterschiedliche Anforderungen an die Usability einer Seite stellen [vgl. Thissen, Schweibenz 2003:63].

Abschließend lässt sich festhalten, dass sich das Internet in seiner Gesamtheit durch die starke Heterogenität von sowohl Nutzerzahl, Zielgruppe als auch Anwendungskontext auszeichnet. Diese Faktoren müssen bei der Usability-Evaluation im Web-Kontext berücksichtigt und Evaluationsmethoden entsprechend angepasst werden.

## 2.2 USER EXPERIENCE

Während der Begriff der Usability die Gebrauchstauglichkeit eines Produktes beziehungsweise einer Website bezeichnet und so eine messbare Größe darstellt, definiert der neuere Begriff der



User Experience das eher subjektive Konzept des Nutzungserlebnisses und der individuellen Einstellung des Nutzers zum und seiner Erfahrung mit dem Produkt [s. Garrett 2003:10f.]. Es ist anzumerken, dass für den Begriff User Experience bisher keine feste Definition vorliegt. Die im Entstehen begriffene ISO-Norm „ISO-CD-9241-210“ definiert User Experience entsprechend als

*„...all aspects of the user's experience when interacting with the product, service, environment or facility. [...] It includes all aspects of usability and desirability of a product, system or service from the user's perspective.“*

[Steward 2008<sup>2</sup>]

Usability und User Experience stellen also nicht zwei disjunkte Konzepte dar, vielmehr wird in der ISO-CD 9241-210 Usability als Teil von User Experience gesehen. Entsprechend setzt sich unter Experten die Einstellung durch, dass die reine Gebrauchstauglichkeit eines Produktes oder einer Anwendung noch kein Garant für ein positives Nutzungserlebnis des Anwenders ist. In Anlehnung an die sogenannte „Zwei-Faktoren-Theorie“ von Herzberg [Herzberg 1959 in Hassenzahl et al. 2008:79] würde eine rein gebrauchstaugliche Anwendung, die es dem Nutzer ermöglicht, seine Ziele unter dem höchstmöglichen Ausmaß an Effektivität und Effizienz zu erfüllen, zunächst einen Zustand der Neutralität, nicht aber des positiven Nutzungserlebnisses beim Anwender hervorrufen. Erst durch das Hinzukommen weiterer Faktoren, die über die reine Usability hinausgehen, kann Zufriedenheit und ein positives Nutzungserlebnis geschaffen werden. Zu nennen wären hier unter anderem Konzepte wie Innovation, Ästhetik, Spaß oder Herausforderung [s. Hassenzahl et al. 2008:78f.] oder die Intuitivität der Bedienung eines Systems [s. Kindsmüller, Mahlke 2007:1f.]. Dieses Modell ist analog übertragbar auf den Web-Kontext.

Bei Hassenzahl et al. (2003) wird zwischen hedonischer (Fähigkeit des Systems, den Nutzer zu fördern und Identität zu kommunizieren) und pragmatischer (Ausmaß der Anwendungstauglichkeit der Anwendung) Qualität unterschieden. Entsprechend kann die pragmatische Qualität als annähernd deckungsgleich mit dem klassischen Begriff der Usability aufgefasst werden. Die Autoren führen im Zusammenhang mit der hedonischen Qualität die zwei Begriffe Stimulation und Identität an. Stimulation bezeichnet hierbei die Art und Weise, auf die der Nutzer durch die Verwendung des Produktes gefördert wird, neue Fähigkeiten erlernt und bestehende Fähigkeiten verbessern kann. Als Identität wird das Ausmaß, in dem sich der Nutzer durch das Produkt in seiner Selbstwahrnehmung bestätigt fühlt und sich in dem Produkt wiederfinden kann, bezeichnet [vgl. Hassenzahl et al. 2003:187f.]. Die letztendlich wahrgenommene Attraktivität des Produkts oder der Anwendung stellt die

*„globale Bewertung auf Basis der wahrgenommenen Qualitäten...“*

[a.a.O.:189]

dar.

---

<sup>2</sup> <http://www.system-concepts.com/articles/usability-articles/2008/usability-or-user-experience-whats-the-difference.html?>

Insofern die Gebrauchstauglichkeit eines Produktes als annähernd äquivalent zur pragmatischen Qualität betrachtet wird, kann User Experience bei den Autoren ebenfalls als übergeordnetes Konzept zu Usability eingeordnet werden und stellt somit nur einen Teil der Nutzungserfahrung dar. In Übertragung dieses Modells auf den Anwendungskontext Internet muss unter Beachtung der vorherig geschilderten Bereitschaft des Nutzers, eine Website bei Negativeindrücken oder Nutzungsproblemen zu verlassen und unter Umständen nicht mehr aufzusuchen, User Experience einen wichtigen Einflussfaktor darstellen, der bei der Erstellung und Evaluation von Webauftritten Beachtung finden sollte. Um einen Nutzer dauerhaft zu binden, muss ein Webprojekt oder ein Internetauftritt zum einen in der Lage sein, sich gegenüber konkurrierenden Angeboten durch eine hohe pragmatische Qualität auszuzeichnen (das heißt, ein hohes Maß an Gebrauchstauglichkeit aufzuweisen), und es ihm zu ermöglichen, sein Nutzungsziel effektiv und effizient zu umzusetzen. Zum anderen muss dem Benutzer der Seite durch die Umsetzung hedonischer Qualitätsaspekte ein memorables und dauerhaft positives Nutzungserlebnis geschaffen werden.

User Experience stellt somit einen zentralen Punkt in der aktuellen Forschung dar. Da zum momentanen Zeitpunkt jedoch noch keine einheitliche Norm zur Definition des Konzeptes vorliegt, soll sich die vorliegende Arbeit schwerpunktmäßig mit der Usability von Websites auseinandersetzen. Mit maßgeblich für diese Entscheidung war vor allem der Umstand, dass bei einer hohen Anzahl von Websites grundlegende Usability-Probleme bestehen. Für den ersten Nutzungseindruck wiegen diese sicherlich schwerer als deren (möglicherweise ebenfalls begrenzte) hedonische Qualität. Wünschenswert wäre entsprechend im ersten Schritt die Behebung grundlegender Nutzungsprobleme und im zweiten, nicht weniger wichtigen Schritt die Optimierung hedonischer Aspekte. Aus diesem Grund soll in der vorliegenden Arbeit versucht werden, grundlegende Usability-Aspekte verschiedener Web-Kontexte sowie ein Instrument zu deren Evaluierung in den Mittelpunkt zu stellen, um so die Grundlage für mögliche weitere hedonische Optimierung zu schaffen.

## 2.3 DIMENSIONEN VON WEB-USABILITY

Nach Beier und von Gizycki (2002) weisen Websites mindestens drei Dimensionen auf, die bei der Festlegung und Spezifikation von Usability-Aspekten berücksichtigt werden müssen: Content (Inhalte der Seite), Design (visuelle Gestaltung) und Struktur (Anordnung der einzelnen Seiten) [s. von Gizycki, Beier 2002:5]. Bei Nielsen (2000) ist eine vergleichbare Kategorisierung unter den Begriffen Content-Design, Page Design und Site-Design zu finden [s. Nielsen 2000:15].

Effektivität, Effizienz und Zufriedenheit müssen für jede dieser Dimensionen einzeln und unterschiedlich definiert werden. In Bezug auf die Inhalte der Website wird die Effektivität und Effizienz eines Systems dadurch bestimmt, dass die Informationen und Inhalte, die der Nutzer benötigt, in einer akzeptablen Zeitspanne (bezogen vor allem auf Ladezeiten der Website) und unter einem vertretbaren Einsatz von Ressourcen (kognitiver Aufwand zum Auffinden der

benötigten Information) gefunden werden können. Zufriedenheit in Relation mit dem Content einer Website kann beispielsweise durch Individualisierung der präsentierten Inhalte erreicht werden.

In Zusammenhang mit dem Design einer Website wird Effektivität durch eine Gestaltung erreicht, die es dem Benutzer erlaubt, gedanklich unterschiedliche Funktionen unterschiedlichen Komponenten zuzuordnen, beispielsweise durch die Anwendung von Gestaltprinzipien oder optischer Metaphern. Wird hierdurch eine einfache und verständliche Orientierung auf der Website erreicht, so liegt Effizienz des Designs vor. Zufriedenheit bezüglich des Designs ist stark zielgruppenabhängig (Geschlecht, Alter) und sollte in Abhängigkeit von den jeweiligen Bedürfnissen der Nutzer umgesetzt werden.

Eine effektive Struktur wird durch eine inhaltlich sinnvolle Anordnung der einzelnen Seiten erreicht, die für den Nutzer genau dann effizient ist, wenn sie für ihn erkenn- und nachvollziehbar ist und zugleich Orientierungshilfe bietet. Zufriedenheit kann wiederum durch Individualisierung (beispielsweise von Navigationspfaden) erreicht werden [s. von Gizycki, Beier 2002:5ff.].

	Effektivität	Effizienz	Zufriedenheit
Content	Auffinden benötigter Inhalte	Akzeptable Zeitspanne zum Auffinden benötigter Inhalte	Individualisierung von Inhalten
Design	Funktionale Gestaltung der Website	Einfache Orientierungsmöglichkeiten, Verständlichkeit	Zielgruppenspezifische Designelemente
Struktur	Sinnvolle Anordnung der einzelnen Seiten	Einfache & eindeutige Navigation möglich	Individualisierung von Navigationspfaden

**Tabelle 1: Dimensionen von Web-Usability**

## 2.4 USABILITY-PROBLEME

Usability-Probleme treten entsprechend immer dann auf, wenn Anforderungen an Effektivität, Effizienz oder Zufriedenheit nicht erfüllt werden können. Um zu definieren, wodurch ein Usability-Problem gekennzeichnet wird, gibt die Literatur unterschiedliche Anhaltspunkte. So liegt bei Sarodnick und Brau ein Usability-Problem vor, wenn

*„ ... Aspekte eines Systems es Nutzern mit hinreichender Domänenenerfahrung unangenehm, ineffizient und beschwerlich oder unmöglich machen, in einem typischen Anwendungskontext die Ziele zu erreichen, für deren Erreichung das System erstellt wurde.“*

Die Autoren unterscheiden hierbei zwischen offenen (auf den ersten Blick ersichtlich) und verdeckten (nur in bestimmten Anwendungsfällen in Erscheinung tretend) Usability-Problemen, eine Klassifikation, die jedoch nicht mit dem Schweregrad der einzelnen Probleme im Zusammenhang stehen muss [vgl. ebd.]. Schweibenz und Thissen halten hierzu in Anlehnung an Jakob Nielsen fest:

*„Generell ist ein Usability-Problem jeder Aspekt einer Benutzerschnittstelle, der die Benutzer voraussichtlich vor Probleme stellen wird[...]. Mögliche Probleme können sich auf verschiedene Aspekte beziehen wie Erlernbarkeit, Grad der Aufgabenerfüllung oder der subjektiven Zufriedenheit.“*

[Schweibenz, Thissen 2003:38]

Es lässt sich also festhalten, dass Usability-Probleme dadurch gekennzeichnet werden, dass der Nutzer an der Erledigung seiner Aufgabe auf jedwede Art oder Weise gehindert wird. Usability-Evaluation wird entsprechend mit dem Ziel vorgenommen, Usability-Probleme aufzudecken und wenn möglich Handlungsoptionen zu deren Behebung oder Verbesserung aufzuzeigen. Wichtig ist hierbei, nicht nur Experten (Personen mit umfangreichen Vorkenntnissen bezüglich des Produkts/der Anwendung oder Vorwissen im Bereich Usability), sondern auch Endnutzer mit einzubeziehen. Beide Personengruppen beurteilen ein Produkt auf Basis unterschiedlichen Hintergrundwissens und auf unterschiedliche Anwendungsarten und können so unter Umständen verschiedene Usability-Probleme aufdecken.

## 2.5 FAZIT

Im vorangegangenen Abschnitt wurde ein kurzer Abriss über die Konzepte Usability, User Experience sowie Dimensionen und Probleme von Usability geliefert. Es lässt sich festhalten, dass sowohl Aspekte der Gebrauchstauglichkeit als auch hedonische Faktoren zur Benutzerzufriedenheit beitragen und bei der Konzeption und Erstellung von Webprojekten Berücksichtigung finden sollten. Von essentieller Wichtigkeit ist vor allem die Beseitigung grundlegender Nutzungshindernisse (Usability-Probleme). Im folgenden Teil sollen entsprechend Methoden der Usability-Evaluation beschrieben und erläutert werden, die sich zum Ziel machen, derartige Probleme aufzudecken und Möglichkeiten zu deren Beseitigung aufzuzeigen. Hierbei sollen sowohl expertenbasierte als auch nutzerorientierte Methoden und Verfahren Berücksichtigung finden.

### 3 METHODEN DER USABILITY-EVALUATION

Um die Gebrauchstauglichkeit und das Nutzungserlebnis von Produkten und Anwendungen beurteilen zu können, müssen beide Konzepte als quantitativ oder qualitativ messbare Größen operationalisiert werden. Eine Evaluation wird nach Schweibenz und Thissen unter anderem durch folgende Eigenschaften gekennzeichnet:

- Ermöglicht eine Bewertung von Handlungsalternativen
- Stellt eine Planungs- und Entscheidungshilfe dar
- Wird mit dem Ziel vorgenommen, etwas zu überprüfen und/oder zu verbessern

[s. Schweibenz, Thissen 2003:38]

Methoden zur Usability-Messung und -Evaluation existieren in umfangreicher Anzahl und lassen sich in einer ersten Klassifikation in benutzerorientierte und expertenorientierte Verfahren unterteilen. Zu unterscheiden ist ebenfalls zwischen formativer (Evaluation während der Entwicklung des Produktes mit dem Ziel iterativer Verbesserung bereits während des Designprozesses) und summativer (Evaluation nach Abschluss des Produktes mit dem Ziel einer Bewertung der Gesamtqualität) Evaluation [s. Sarodnick, Brau 2006:20].

Im Folgenden soll nur ein kurzer Überblick über Methoden zur Usability-Evaluation gegeben werden, für eine ausführliche Beschreibung siehe z.B. Sarodnick, Brau (2006).

#### 3.1 EXPERTENORIENTIERTE METHODEN DER USABILITY-EVALUATION

Expertenorientierte oder auch analytische Methoden werden mit dem Ziel angewandt, unter Einsatz von Experten als sogenannten „Ersatz-Usern“ [Schweibenz, Thissen 2003:43] die Gebrauchstauglichkeit eines Produktes oder einer Anwendung zu ermitteln, wobei die Beurteilung anhand von festgelegten Richtlinien (Heuristiken) und des jeweiligen Expertenwissens erfolgt

[s. Sarodnick, Brau 2006:113]. Die Interaktion des Nutzers mit dem Gegenstand der Evaluation wird hierbei durch einen Experten simuliert. Die bekanntesten Beispiele für analytische Methoden der Usability-Evaluation sind die Heuristische Evaluation (Exploration des Systems, Aufdecken und Protokoll von Verstößen gegen einen Satz von Usability-Prinzipien (Heuristiken)) und der sogenannte Cognitive Walkthrough (Entwicklung einer multistep-Ideallösung zur Erfüllung typischer Aufgaben mit dem System und anschließende Analyse der verschiedenen Schritte auf mögliche Probleme für den Nutzer unter Berücksichtigung zielgruppenspezifischer Aspekte) [a.a.O.:134ff.]. Analytische Methoden sind – bedingt durch ihren geringeren Aufwand im Vergleich mit empirischen Methoden – vor allem in frühen Entwicklungsphasen einsetzbar und können in späteren Phasen als Ergänzung zu benutzerorientierten Verfahren dienen [vgl. a.a.O.:113 f.].

### 3.2 BENUTZERORIENTIERTE METHODEN DER USABILITY-EVALUATION

Bei den benutzerorientierten oder auch empirischen Verfahren werden die Informationen aus Erhebung und Analyse der Gebrauchstauglichkeit direkt von den Nutzern gewonnen und müssen im Anschluss interpretiert werden: Sie stellen zum Teil subjektive Eindrücke dar, die unter Umständen von Erfahrungsstand und Persönlichkeit der befragten Person abhängen können. Bekannte Beispiele für benutzerorientierte Methoden der Usability-Evaluation sind Usability Lab-Tests (Test des Produktes durch Endnutzer im Usability-Labor unter Beobachtung von Experten; Dokumentation von tatsächlich erfolgtem Verhalten) und die Nutzerbefragung durch qualitative Interviews und Fragebögen (Dokumentation von Aussagen über potentiell Verhalten) [s. Schweibenz, Thissen 2003:43]. Empirische Methoden zur Aufdeckung von Usability-Problemen weisen den Nachteil eines hohen Aufwands (Zeit, Kosten) auf, erbringen jedoch (im Gegensatz zur Expertenevaluation) gleichzeitig den Vorteil von echten Nutzermeinungen und Einschätzungen. Vor allem in der summativen Evaluation von Produkten und Anwendungen sollten benutzerorientierte Verfahren zum Einsatz kommen, um die Einsetzbarkeit unter und die Akzeptanz durch die Endanwender zu gewährleisten [vgl. a.a.O.:118f.].

### 3.3 FAZIT

Im vorangegangenen Abschnitt wurden kurz sowohl experten- als auch nutzerzentrierte Methoden der Usability-Evaluation vorgestellt. Beide Verfahren werden mit dem Ziel durchgeführt, potentielle Probleme, die die Gebrauchstauglichkeit einer Anwendung beeinträchtigen können, aufzudecken und eventuelle Lösungsmöglichkeiten zu deren Behebung aufzuzeigen. Einen wichtigen Bestandteil der empirischen Methoden stellt die Nutzerbefragung dar. Im folgenden Abschnitt soll näher auf das Verfahren der Usability-Evaluation mithilfe von Fragebögen eingegangen, sowie Vor- und Nachteile des Verfahrens herausgestellt werden.

## 4 EMPIRISCHE USABILITY-EVALUATION MIT FRAGEBÖGEN

Die empirische Sammlung von Nutzerdaten über potentiell oder tatsächlich erfolgtes Verhalten mit Fragebögen stellt, wie im vorherigen Abschnitt beschrieben, neben qualitativen Interviews oder Usability-Lab-Tests eine weit verbreitete Form der Datenerhebung dar. Hierbei werden die Nutzer mittels vollstandardisierter (vorgegebene Itemformulierung und Antwortkategorien, geschlossene Fragen), teilstandardisierter (zusätzliche Freitextantworten möglich) oder nichtstandardisierter (reine Freitextantworten) Fragebögen um eine subjektive Einschätzung von Merkmalsausprägungen bestimmter Produkt- oder Anwendungseigenschaften gebeten. Zu beachten ist hierbei, dass die Aussagen der Benutzer Einflussfaktoren wie sozial erwünschtem

Verhalten<sup>3</sup> oder Rationalisierung des eigenen Verhaltens<sup>4</sup> unterliegen und somit verzerrt werden können. Auch stimmen Eigen- und Fremdwahrnehmung des Nutzungsverhaltens nicht immer überein. Befragungsmethoden können und sollen daher eine gute Ergänzung zu formativen und summativen Methoden im Usability-Lab darstellen, ihre alleinige Anwendung zur Usability-Evaluation kann jedoch zu verzerrten Ergebnissen führen [vgl. Schweibenz, Thissen 2003:119].

## CHECKLISTEN, QUESTIONNAIRES, SURVEYS

Im Kontext von Nutzerbefragungen zur Usability-Evaluation lassen sich in Anlehnung an Kirakowski (1994) mehrere Typen von Benutzerbefragungen unterscheiden: Checklisten, Surveys und sogenannten Questionnaires. Sie unterscheiden sich in Anwendungskontext, Itemformulierung und Standardisierung und sind deshalb für jeweils unterschiedliche Anwendungsbereiche geeignet.

Von Benutzerbefragungen mittels Skalen-Fragebögen, die die Einstellung der Befragten zu einem Thema oder einem Produkt wiedergeben (Questionnaires) sind vor allem die Checkliste und das Survey abzugrenzen. Bei Checklisten-Erhebungen wird mit einem Satz von Eigenschaften, die das System wünschenswerterweise besitzen sollte, deren Vorhandensein oder Abwesenheit überprüft. Checklisten werden vor allem von Experten im Rahmen heuristischer Evaluationen mit Experten eingesetzt, können aber durchaus auch für die Erhebung von Endnutzerdaten verwendet werden. Surveys werden vor allem zur Erhebung objektiv beurteilbarer Eigenschaften eines (Teil-)Systems eingesetzt. Sie grenzen sich von Einstellungsfragebögen vor allem durch die Eigenschaft der objektiven Messbarkeit der zu erhebenden Attribute ab [Kirakowski 1994<sup>5</sup>]. Die in der folgenden Arbeit vorliegenden Fragebögen sind entsprechend den sog. Questionnaires (dt. Fragebögen) zuzuordnen.

### 4.1 FRAGEBOGENAUFBAU

Fragebögen zur Erhebung quantitativer und qualitativer Daten setzen sich aus einem bestimmten Satz an Items (Fragen/Aussagen) zusammen, der wiederum in Subskalen (Unterthemen) unterteilt sein kann. Die jeweiligen Items erheben jeweils die Ausprägung eines Merkmals beim Nutzer, beispielsweise seine subjektive Einschätzung und/oder seine Einstellung zu einem bestimmten Objekt oder Sachverhalt. Eine vorhergehende ausreichende Auseinandersetzung des Nutzers mit dem zu evaluierenden Objekt ist also eine notwendige Bedingung.

Fragebogenitems können je nach Antwortoption in Multiple-Choice (eine aus mehreren Antworten möglich), Mehrfachauswahl (mehrere Antworten möglich), Ratingskalen (bipolare Einschätzung) oder offene Fragen (mit Freitextantwort) untergliedert werden. Wichtig bei der

---

<sup>3</sup> Der Befragte gibt die Antwort, von der er annimmt, dass sie gesellschaftlich sozial erwünscht wäre.

<sup>4</sup> Unbewusste Rechtfertigung des eigenen Verhaltens [s. Myers 2004:572]

<sup>5</sup> <http://sumi.ucc.ie/sumipapp.html>

Fragebogenkonstruktion sind vor allem die sogenannten Gütekriterien Objektivität (Ergebnisse unabhängig vom Evaluator?), Reliabilität (gleiche Ergebnisse bei Wiederholung der Erhebung?) und Validität (wird das zu messende Konstrukt richtig erfasst?) sowie eine Standardisierung mit einer ausreichend großen Stichprobe (siehe Abschnitt 5) [vgl. Bortz, Döring 2006:213 ff.]. Vollstandardisierte Fragebögen haben vor allem den Vorteil einer leichteren Auswertung, da eine rein numerische und dadurch leicht zu automatisierende Erfassung der Antwortkategorien möglich ist. Sie sind so vor allem für die Web-Anwendung in großem Maßstab geeignet [vgl. Sarodnick, Brau 2006:172f.].

## 4.2 COMPUTERVERMITTELTE BEFRAGUNGEN

Bedingt durch das rasche Wachstum des Internet und die Technologisierung hat die Verbreitung von Nutzerbefragungen per Online-Fragebögen immer mehr zugenommen. Geschätzt wird, dass mittlerweile ein Drittel aller durchgeführten Nutzerbefragungen über das Medium Internet durchgeführt werden [s. Jakob et al. 2009:9]. Vor allem die hohe Zahl potentiell erreichbarer Teilnehmer sowie eine geringere Fehlerquote bei Eingabe und Auswertung (Echtzeitkontrollen von fehlerhaften oder fehlenden Eingaben möglich) haben, zusammen mit einer hohen wahrgenommenen Anonymität der Befragten dazu geführt, dass Nutzerbefragungen nicht mehr auf postalischem Weg, sondern direkt im Internet durchgeführt werden [vgl. Sarodnick, Brau 2006:172f.]. Es werden hierbei Kosten für Material (Anschriften per Brief, Papierbögen) gespart und dem Nutzer kann – in begrenztem Ausmaße – interaktiv Hilfestellung gegeben werden. Auch Voreingenommenheit kann sowohl bei Befragter als auch Befragten so vermieden [vgl. Lorenzo-Seva, Ferrando 2005:193] und die Motivation des Befragten durch den Einsatz multimedialer Elemente gesteigert werden [vgl. Jakob et al. 2009:26f.].

Neben den genannten Vorteilen bringt die Methode der Online-Befragung jedoch die Nachteile einer Verzerrung der Stichprobe (wenig Kontrolle über die Zusammensetzung der Teilnehmergruppe in Kombination mit einem erhöhten Anteil internetaffiner und potentiell jüngerer Nutzer) und eine erhöhte Abbruchbereitschaft der Teilnehmer durch die Unverbindlichkeit des Mediums Internet mit sich [vgl. Sarodnick, Brau 2006:172]. Bevölkerungsrepräsentative Studien können und sollten demnach nicht durch Online-Befragungen erhoben werden, solange nicht durch einschlägige Maßnahmen (Restriktion des Teilnehmerkreises, Verwendung eines Online-Panels<sup>6</sup>) die Repräsentativität der Zielgruppe sichergestellt werden kann [vgl. Jakob et al. 2009:23ff.].

---

<sup>6</sup> Pool registrierter Teilnehmer, die repräsentativ die Gesamtbevölkerung abbilden



#### 4.3 FRAGEBOGENKONSTRUKTION UND GESTALTRICHTLINIEN ZUR ONLINE-ERSTELLUNG

Obwohl mittlerweile Online-Befragungen zum Teil ausschließlich für das Medium Internet entwickelt werden, sind viele Fragebögen für Online-Nutzerbefragungen direkte Übertragungen aus sog. „Pen and Paper“- Fragebögen [vgl. Lorenzo-Seva, Ferrando 2005:193], was vor allem durch deren Aufbau und Länge zu einer Erhöhung der Nonresponse-Quote (Abbruchrate) führen kann [vgl. Kaczmirek 2008: 40f.].

Zur Gestaltung von Online-Fragebögen können folgende Richtsätze genannt werden:

1. Beachtung der technischen Anforderungen
2. Formulierung der Fragen nach den Regeln der empirischen Sozialforschung
3. Etablierung einer glaubwürdigen Kommunikation mit den Antwortenden
4. Erzeugung von Awareness (in dt. Literatur „Gewärtigkeit“) und Aufmerksamkeit
5. Verwendung eines ansprechenden Designs und Umsetzung von Usability-Kriterien
6. Hervorhebung zentraler Textpassagen
7. Vermeidung von Matrix-Fragen
8. Begrenzung der Länge des Fragebogens
9. Einsatz von Filter-Fragen
10. Durchführung eines Vortests

[s. Gräf 2002:78]

Maurer und Jandura (2009) weisen in diesem Zusammenhang ebenfalls darauf hin, Texte im Sinne der Benutzerfreundlichkeit kurz zu halten, um die Abbruchwahrscheinlichkeit zu verringern, sowie eine Nummerierung der Items vorzunehmen und eindeutige Kontraste zur Fragenkennzeichnung zu verwenden. Online-Fragebögen sollten im Design an Papierfragebögen angepasst werden (dunkle Schrift auf hellem Hintergrund). Fehlermeldungen sollten verständlich gehalten sein. Fortschritts- oder Verlaufsanzeigen wirken sich positiv auf die Nonresponse-Quote aus und tragen zur Nutzerorientierung bei. Ebenfalls muss auf Browserkompatibilität bei der Erstellung des Fragebogens geachtet werden. Bei der Skalenwahl sollte vor allem auf das Erscheinen aller Antwortkategorien auf dem Bildschirm geachtet werden sowie auf gleichmäßigen Abstand zwischen den Optionsfeldern, deren Anzahl nicht zu groß sein sollte. Vor allem die Möglichkeit von Verzweigungsfiltern innerhalb des Fragebogens sollte Gebrauch finden [vgl. Maurer, Jandura 2009:67ff.]. Um eine ansprechende Auswertung der erhobenen Nutzerdaten zu gewährleisten, können die Daten beispielsweise in einem sogenannten grafischen Profil, auf dem Mittelwerte und Abweichungen visualisiert werden, abgebildet werden. Auch Histogramme (Säulendiagramme) sind an dieser Stelle häufig in Gebrauch [vgl. Sarodnick, Brau 2006:172f.]. Vor- und Nachteile des zu evaluierenden Systems werden bei entsprechender Polung der Itemantworten in der Auswertung so auf den ersten Blick erkennbar.

#### 4.4 TEILNEHMERKREIS & STICHPROBENWAHL

Online-Befragungen können in Foren, per E-Mail oder auf festgelegten Websites durchgeführt werden, und können durch die Art der Stichprobenziehung in probabilistische und nicht-probabilistische Verfahren unterteilt werden [s. Bortz, Döring 2006:402f.]. Probabilistische Verfahren zeichnen sich dadurch aus, dass die Befragung mit einer statistisch zufälligen Auswahl arbeitet (Festgelegter Teilnehmerkreis, z.B. Online-Panel) wohingegen an nichtprobabilistische Verfahren eine willkürliche Auswahl an Personen teilnimmt (Freischaltung der Online-Befragung für alle potentiellen Teilnehmer) [vgl. Maurer, Jandura 2009:62]. Auf die Nachteile nichtprobabilistischer Verfahren wurde bereits in Abschnitt 4.2 hingewiesen.

Durch ihren inhärent höheren Anteil internetaffiner Nutzer ist das Medium der Online-Befragung vor allem in der Evaluation von Online-Auftritten und Web-Projekten eine beliebte und sinnvolle Methode, um empirische Erhebungen von Nutzerdaten durchzuführen. Zur Online-Evaluation von Websites steht bereits eine Reihe standardisierter und den allgemeinen Gütekriterien (Objektivität, Validität und Reliabilität) entsprechenden Fragebögen zur Verfügung, die jedoch meist allgemeiner Natur sind und nicht auf einen jeweiligen Nutzungskontext zugeschnitten. Wie in Abschnitt 2.1 jedoch bereits angeführt, kann Usability im Web nur schwer pauschal beurteilt werden, sondern erfordert stets den Einbezug des Nutzungskontextes. Im folgenden Abschnitt sollen daher zunächst bestehende Fragebögen zur Evaluation von Webprojekten und Online-Auftritten vorgestellt und analysiert werden, um im Anschluss deren Tauglichkeit für den Einsatz in einem spezifischen Nutzungskontext zu bewerten.

#### 4.5 FAZIT

Abschließend lässt sich festhalten, dass Fragebögen in Ergänzung zu analytischen und beobachtenden empirischen Methoden eine gute und unverzichtbare Methode in der Usability-Evaluation darstellen. Online-Benutzerbefragungen erweisen sich vor allem Einsatz bei der Evaluation von Webauftritten und -projekten als nützliches und sinnvolles Verfahren.

Bei Konstruktion und Einsatz von Fragebögen müssen jedoch bestimmte grundlegende Prinzipien beachtet werden, um aussagekräftige Daten erheben zu können. Daher soll im folgenden Abschnitt ein kurzer Überblick über das psychologische Teilgebiet der Testtheorie sowie der Fragebogen-konstruktion geliefert und der Einsatz von Online-Fragebögen unter wissenschaftlichen und praktischen Gesichtspunkten beurteilt werden.

In der sogenannten Testtheorie über psychologische Tests <sup>7</sup> werden im Allgemeinen psychologische Konstrukte (Ausprägungen von Eigenschaften, Persönlichkeitsmerkmalen bei Personen) mittels Fragebögen, Interviews oder Beobachtungen erfasst, die für die Beteiligten möglichst nicht willentlich verfälschbar sein sollten.

Die Klassische Testtheorie, auf der ca. 95% aller Tests basieren [s. Rost 1999:140 in Bühner 2006:21],

*„...beschäftigt sich mit der Frage, wie das Testverhalten einer oder mehrerer Personen von einem zu erfassenden psychischen Merkmal abhängt“,*

[Bühner 2006:22]

die Probabilistische Testtheorie hingegen

*„...befasst sich mit den unterschiedlichen Bestandteilen von Messwerten (wahrer Wert + Fehler).“*  
[ebd.].

Standardisierte Tests und Fragebögen, die in Anlehnung an testtheoretische Kriterien entworfen werden, werden im Allgemeinen als psychometrischer Test oder Skala bezeichnet. Ein psychometrischer Test ist demnach ein

*„...wissenschaftliches Routineverfahren zur Untersuchung eines oder mehrerer empirisch abgrenzbarer Persönlichkeitsmerkmale mit dem Ziel einer möglichst quantitativen Aussage über den relativen Grad der individuellen Merkmalsausprägung.“*

[a.a.O.:23]

Die Benutzerbefragung per Fragebogen und die daraus resultierende Erhebung empirischer Daten ist in der Sozialforschung den quantitativen Methoden zur Datenerhebung zuzuordnen. Sie

*„...erfordert eine hohe Strukturierbarkeit der zu erfassenden Inhalte und verzichtet auf steuernde Eingriffe eines Interviewers.“*

[Bortz, Döring 2006:252],

zeichnet sich also durch eine wesentlich geringere Kontrolle des Befragers (Evaluators) über die Befragungssituation, Verständnisprobleme des Befragten oder dessen Umgehen mit dem Fragebogen (Auslassen von Items etc.) aus. Fragebögen können entweder Einstellungen und Meinungen der befragten Personen erfassen oder deren konkrete Verhaltensweisen, Fragebögen ersteren Typs werden den psychometrischen Persönlichkeitstests zugeordnet. Im Kontext von Usability-Evaluationen handeln sich die in der vorliegenden Arbeit zu erstellenden Itemsets um

---

<sup>7</sup> abzugrenzen von der Testtheorie über statistische Schlüsse, die zum Test einer Hypothese oder Person verwendet wird.

Persönlichkeitstests vom Erhebungstyp der Querschnittsdiagnose, da hier ein momentan gültiger Zustand erhoben und abgebildet wird [vgl. Bühner 2006:14f.].

## 5.1 FRAGEBOGENKONSTRUKTION

Zu beachten bei der Fragebogenkonstruktion ist eine möglichst breite Abdeckung des Gegenstandsbereichs und seiner Inhalte durch die ausgewählten Items, eine themenspezifische Klassifikation der Items in Untergruppen sowie die zielgruppenspezifische Anpassung der sprachlichen Formulierungen der Fragebogenitems [vgl. Bortz, Döring 2006:252]. In der vorliegenden Arbeit handelt es sich um den Gegenstandsbereich der Usability von bestimmten Website-Typen. Zur Abdeckung des gewählten Bereichs durch die Fragebogenitems siehe Abschnitt 7.4.

Zur Formulierung der Items lässt sich festhalten, dass hinsichtlich einer erleichterten Auswertung vorgegebene Antwortkategorien, wie sie in vollstandardisierte Fragebögen vorkommen, offenen Antwortmöglichkeiten vorzuziehen sind. Einstellungen und Meinungen können durch Aussagen (sog. Statements<sup>8</sup>, beispielsweise Likert-Skalen) erhoben werden, da sie den Befragten eindeutige Stellungnahmen über Zustimmung oder Ablehnung zu diesen abverlangen. Konkrete Sachverhalte hingegen werden vorzugsweise durch direkte Fragen<sup>9</sup> ermittelt, erfordern jedoch größere Vorarbeit bezüglich der vorgegebenen Antwortkategorien.

Weiterhin ist darauf zu achten, Fragebogenitems möglichst neutral zu formulieren (beziehungsweise unterschiedliche Wertigkeiten in der Formulierung vorzunehmen, sollte zum gleichen Sachverhalt mehr als ein Item verwendet werden<sup>10</sup>) um eine Verfälschung der Antworten zu vermeiden. Auf eine eindeutige und verständliche Ausdrucksweise sollte ebenfalls geachtet werden. Das Ende des Fragebogens sollte aus leicht zu beantwortenden Items bestehen.

Zu vermeiden sind in jedem Fall:

- Verabsolutierende oder hypothetische Aussagen
- Übermäßig lange Fragen und/oder Aussagen
- Suggestivfragen
- Konjunkte Antwortkategorien

[vgl. Porst 2000 in Bortz, Döring 2006:252]

---

<sup>8</sup> Beispiel: „Die Website bietet eine ausreichende Hilfefunktion.“

<sup>9</sup> Beispiel: „Wie viel Zeit verbringen sie durchschnittlich pro Tag im Internet?“

<sup>10</sup> Beispiel: „Ich bin der Ansicht, dass sich Computerspiele negativ auf das Sozialverhalten auswirken“ vs. „Ich bin der Ansicht, Computerspiele haben keinen Einfluss auf das Sozialverhalten“

## 5.2 HAUPT- UND NEBENGÜTEKRITERIEN

Zur Beurteilung der Güte eines psychometrischen Tests dienen die sogenannten Gütekriterien. Diese werden in Hauptgütekriterien (Objektivität, Reliabilität und Validität) sowie Nebengütekriterien (Normierung, Vergleichbarkeit, Ökonomie, Nützlichkeit) unterteilt.

### 5.2.1 HAUPTGÜTEKRITERIEN

#### 5.2.1.1 OBJEKTIVITÄT

Unter der Objektivität eines psychometrischen Tests versteht man das Ausmaß der Abhängigkeit der Testergebnisse vom Testleiter (Untersuchenden). Zu unterscheiden sind hierbei

- Durchführungsobjektivität (festgelegter Versuchsablauf und gleichartige Testinstruktionen bei jeder Durchführung des Tests)
- Auswertungsobjektivität (festgelegte Auswertungsvorschriften, Kontrolle über die Auswertenden zur Sicherstellung der selben Testergebnisse pro Proband unabhängig vom Auswertenden)
- Interpretationsobjektivität (gleiche Interpretation der Testergebnisse unabhängig vom Auswertenden)

[s. Bühner 2006:28f.]

#### 5.2.1.2 RELIABILITÄT

Unter der Reliabilität eines Tests ist der Grad der Messgenauigkeit und entsprechend der Reproduzierbarkeit der Messergebnisse zu verstehen. Es wird ermittelt, inwiefern der Test bei erneuter Durchführung dieselben Ergebnisse liefert. Reliabilität kann also als Korrelationswert <sup>11</sup> zweier Resultate desselben Tests definiert werden [s. a.a.O.:117]. Unterschieden werden hierbei

- Halbierungsreliabilität (Korrelation der Ergebnisse zweier Hälften desselben Tests)
- Retest-Reliabilität (Korrelation der Testleistungen bei erneuter Durchführung des Tests unter denselben Bedingungen und mit denselben Personen nach einem angemessenen Zeitraum)
- Paralleltestreliabilität (Korrelation zwischen zwei Tests bei Messung desselben Merkmals mittels unterschiedlicher Items)

[s. a.a.O.:29]

Wenn es aus Zeit- oder Umfangsgründen nicht sinnvoll oder möglich ist, die oben genannten Verfahren zur Reliabilitätsmessung durchzuführen, ist es möglich, die interne Konsistenz der Skala

<sup>11</sup> Grad der Übereinstimmung

zu schätzen. Hierbei wird jedes Item wie ein eigenständiger Test behandelt, und die Korrelationen der einzelnen Itemergebnisse ermittelt. Bei Tests mit stark heterogenen Aufgaben ist dieses Verfahren jedoch nicht sinnvoll, da keine aussagekräftigen Ergebnisse erzielt werden können [vgl. a.a.O.:118].

---

### 5.2.1.3 VALIDITÄT

Unter dem Konzept der Validität eines psychometrischen Tests wird verstanden, dass der Test misst, was er messen soll. Zu unterscheiden sind hierbei

- Inhaltsvalidität (der Test oder das Testitem erfasst das zu bestimmende Merkmal genau, die Menge aller Items bildet das interessierende Merkmal erschöpfend und genau ab)
- Kriteriumsvalidität (Korrelation der Testleistung mit äußeren Kriterien oder Kennwerten für dasselbe Konstrukt<sup>12</sup>)
- Konstruktvalidität (Korrelation des Tests mit konstruktverwandten Tests<sup>13</sup>)

[s. a.a.O.:29]

Wichtig an dieser Stelle ist, dass sich die Hauptgütekriterien interdependent verhalten, dies bedeutet mangelnde Objektivität wirkt sich mit hoher Wahrscheinlichkeit auf Reliabilität und Validität aus und umgekehrt. Bei der Konstruktion eines Tests ist also auf die Erfüllung aller Hauptgütekriterien in ausreichendem Maße zu achten [a.a.O.:30ff.].

Inwieweit die Gütekriterien der allgemeinen Testtheorie zur Bewertung der in der vorliegenden Arbeit anzufertigen Fragebogenitems geeignet sind, soll in Abschnitt 10 geklärt werden.

---

### 5.2.2 NEBENGÜTEKRITERIEN

#### NORMIERUNG

Das Ausmaß der Normierung eines Tests sagt aus, inwiefern (zielgruppenspezifische) Referenzwerte aus Normstichproben für die Testergebnisse vorliegen, also eine Einordnung und entsprechende Bewertung vorgenommen werden kann. Über Normstichproben sollten Angaben bezüglich Repräsentativität, Entstehung der Stichprobe und Testbedingungen vorliegen.

#### VERGLEICHBARKEIT

---

<sup>12</sup> z.B. einem zuvor ermittelten Kennwert. Beispiel: Korrelation von Intelligenztest und Schulnote.

<sup>13</sup> Das Testresultat sollte mit dem Ergebnis eines unabhängigen, zum gleichen Zeitpunkt durchgeführten anderen Tests, der denselben oder einen ähnlichen Sachverhalt misst, positiv korrelieren.

Das Kriterium der Vergleichbarkeit für einen psychometrischen Test ist dann erfüllt, wenn andere Tests, die denselben Gültigkeitsbereich erfassen, vorliegen. Personen, die in zwei vergleichbaren Tests befragt werden, sollten entsprechend ähnliche Ergebnisse erzielen

## ÖKONOMIE

Ein psychometrischer Test ist genau dann ökonomisch, wenn die Durchführungszeit kurz, der Materialverbrauch gering, eine einfache Durchführung (auch als Gruppentest) möglich und die Auswertung in angemessener Zeit und mit angemessenem Aufwand vorzunehmen ist.

## NÜTZLICHKEIT

Die Nützlichkeit eines psychometrischen Tests gibt eine Aussage darüber, inwieweit das zu erhebende Merkmal eine Notwendigkeit bezüglich der Erhebung aufweist und die Erhebung nicht redundant ist.

[a.a.O.:34]

Die Bewertung der erstellten Itemsets unter den Gesichtspunkten der Nebengütekriterien soll ebenfalls in Abschnitt 10 durchgeführt werden.

### 5.3 EVALUIERUNG DER METHODIK DES ONLINE – FRAGEBOGENVERFAHRENS

Evaluationsmethoden werden im Allgemeinen entsprechend der Anzahl an Usability-Problemen, die durch sie im Vergleich mit anderen Verfahren aufgedeckt werden können, beurteilt. Fragebogenverfahren als Mittel der Evaluation von Nutzerzufriedenheit im Web finden mittlerweile ein immer größeres Anwendungsfeld. Sarodnick und Brau stellen fest, dass

*„...das Evaluationsanliegen, die Art des Systems sowie andere Faktoren der Durchführbarkeit von Evaluationen einen maßgeblichen Einfluss darauf haben, welche der Verfahren sinnvoll sein können.“*

[Sarodnick, Brau 2003:181]

Sarodnick und Brau zeigen, dass Usability Tests im Vergleich mit anderen (expertenbasierten) Evaluationsmethoden in der Literatur unterschiedlich gut abschneiden und sich keine eindeutige Tendenz erkennen lässt, welches die absolut bessere Methode darstellt. Usability Tests weisen bei Fu et al. (2002) Stärken im Bereich der Wissensebene (mentale Modelle, Lernförderlichkeit) auf, wohingegen die heuristische Evaluation im Bereich Gewohnheiten und Fertigkeiten (Wahrnehmung von bspw. Website-Layouts) sowie Regeln (Konsistenz) besser abschneidet [Fu et al. 2002 in Sarodnick, Brau 2003:181ff.]. Wiederum andere Studien<sup>14</sup> zeigen, dass mit Hilfe von

---

<sup>14</sup> Desurvire (1994), Bowen (1994) in Sarodnick, Brau (2003)

Usability-Tests eine höhere Anzahl an Usability-Problemen aufgedeckt werden kann. Cognitive Walkthroughs schneiden durchschnittlich schlechter ab als heuristische Evaluation und auch Usability-Tests [s. Sarodnick, Brau 2003:183].

Es kann also festgehalten werden, dass experten- und nutzerbasierte Methoden unterschiedliche Ansätze darstellen, und sich die resultierenden Informationen entsprechend unterscheiden, was einen direkten Vergleich erschwert. Wenn möglich, sollten beide Verfahren in Kombination eingesetzt werden, um ein möglichst vollständiges Bild von der Gebrauchstauglichkeit einer Website zu vermitteln.

Bei der Bewertung von Websites kann, bedingt durch die große Anzahl potentieller Nutzer meist nicht auf einen festen Personenkreis, der für Umfragen zur Verfügung steht, zurückgegriffen werden. Deshalb steht fest, dass elektronische Fragebogenverfahren aufgrund ihrer kostengünstigen und flexiblen Einsetzbarkeit ein geeignetes Mittel zur Evaluation von Websites darstellen, da sich problemlos auch größeren Gruppen von (zuvor unbekannten) Personen ad hoc befragen lassen. Bezogen auf den Anwendungskontext von [www.usability-toolkit.de](http://www.usability-toolkit.de), stellen Nutzerbefragungen ein geeignetes Instrument zur Beurteilung von Webauftreten und -projekten der Zielgruppe (kleine und mittelständische Unternehmen) dar. Da nicht vorausgesetzt werden kann, dass bei den Mitgliedern der Zielgruppe Heuristische Evaluationen oder Cognitive Walkthroughs mithilfe von Usability-Experten durchgeführt werden können, sind Fragebogenverfahren ein für die Zielgruppe angemessenes und für die zu erhebenden Information (Nutzerzufriedenheit) geeignetes Verfahren, um Evaluationen durchzuführen [a.a.O.:183ff.].

Kriterien, die von Sarodnick und Brau vorgeschlagen werden, um einzelne Evaluationsmethoden objektiv beurteilen zu können, sind zum einen die sogenannte „Praktische Relevanz“ und zum anderen die sog. „Wissenschaftliche Relevanz“. Diese sollen im Folgenden auf die Methode des Fragebogenverfahrens im Webkontext übertragen und somit evaluiert werden.

---

### 5.3.1 PRAKTISCHE RELEVANZ

Bei der Praktischen Relevanz stellt sich vor allem die Frage der Produktivität (Anzahl der identifizierten Probleme), des Aufwandes (materiell, zeitlich), des Vorwissens der Evaluatoren sowie dem Detaillierungsgrad und der Flexibilität des Verfahrens. Fragebogenverfahren zur Online-Erhebungen der Nutzerzufriedenheit zeichnen sich als Usability-Tests dadurch aus, dass sie vor allem bei Aussageskalen (im Gegensatz zu semantischen Differentialen) einen Großteil der vorhandenen Usability-Probleme durch Nutzerbewertungen und –meinungen indizieren können, auch Akzeptanzprobleme können hier unter Umständen aufgezeigt werden. Der materielle Aufwand ist im Vergleich zu herkömmlichen „Pen- and Paper“- Fragebogenverfahren gering (keine Kosten für Papier, etc.) und erfordert lediglich die Erstellung durch einen Programmierer und/oder Webdesigner. Der zeitliche Aufwand für die Evaluatoren ist ebenfalls gering, da Online-



Fragebögen entweder direkt im Netz zugänglich sind oder an die zu befragenden Probanden verschickt werden können. Der zeitliche Aufwand für Konzeption und Vorbereitung entspricht dem eines herkömmlichen Fragebogens, kann also – sofern nicht auf bereits bestehende, standardisierte und getestete Verfahren zurückgegriffen wird, unter Umständen einige Zeit beanspruchen. Entsprechend müssen die Konzeptionierenden über Expertenwissen im Bereich Fragebogenkonstruktion verfügen, bei den durchführenden Evaluatoren ist dies jedoch nicht zwingend erforderlich. Der Detaillierungsgrad der Ergebnisse kann von Fragebogen zu Fragebogen variieren. Sind Freitextantworten innerhalb von teilstandardisierten Fragebögen möglich, liegt eine detaillierte Beschreibung der festgestellten Usability-Probleme vor. Vollstandardisierte Fragebögen verzichten auf diese Option, erfordern dafür jedoch einen geringeren zeitlichen Aufwand in der Auswertung und können zumindest Hinweise auf bestehende Usability-Probleme liefern. Nachteilig wirkt sich lediglich aus, dass vorgegebene Antwortkategorien den Nutzer hinsichtlich seines Urteils einschränken und sich die eigene Entwicklung von Fragebögen als sehr aufwändig erweist. Vor allem bei Online-Befragungen ist eine erhöhte Abbruchquote erzeugt durch die Unverbindlichkeit des Online-Mediums und/oder Fehler im Design sowie die Verzerrung der Repräsentativität der Zielgruppe (internetaffine Personen sind stärker vertreten, als empirisch repräsentativ ist) ein weiterer Nachteil.

Fragebogenverfahren zeichnen sich im Allgemeinen durch eine mittlere Flexibilität bezüglich der Anpassbarkeit an unterschiedliche Kontexte aus. Ein simpler Austausch ist, bedingt durch die Korrelationen der Items und ihre Auswirkungen auf die Reliabilität des Bogens, nur begrenzt möglich. So erfordert eine Anpassung an unterschiedliche Kontexte unter Umständen eine Rekonstruktion des Fragebogens. Subskalen können jedoch gestrichen oder hinzugefügt werden, so dass eine Adaption auf unterschiedliche Projekte oder Teilaspekte innerhalb desselben Kontextes durchaus problemlos möglich ist [vgl. Sarodnick, Brau 2003:184ff.].

Nachfolgend sind die Kriterien der praktischen Relevanz von Fragebogenverfahren schematisch dargestellt:

Faktoren	Erklärung	Beurteilung von Fragebogenerfahren
Produktivität	Anzahl und Schwere der aufgedeckten Usability-Probleme	Globale Bewertungen; Hinweise auf Nutzungsprobleme bei Aussageskalen; Aufdeckung von Akzeptanzproblemen
Materieller Aufwand	Kosten, Material	Im Vergleich zu traditionellen Pen-and-Paper-Fragebögen gering
Zeitlicher Aufwand	Benötigte Zeit	Gering, da direkt für Nutzer im Web oder per Email zugänglich
Notwendige Qualifikation der Evaluatoren	Expertenwissen notwendig?	Expertenwissen bei der Konstruktion eines Fragebogens erforderlich, wenig Expertenwissen für Erfassung/Auswertung nötig
Detaillierungsgrad	Detaillierungsgrad der Beschreibung des Usability-Problems	Abhängig vom Standardisierungsgrad des Fragebogens
Flexibilität	Adaptierbarkeit des Verfahrens auf unterschiedliche Kontexte	Hoch innerhalb desselben Kontextes, geringer bei stark variierenden Kontexten

**Tabelle 2: Praktische Relevanz von Fragebogenverfahren in der Usability-Evaluation**

### 5.3.2 WISSENSCHAFTLICHE RELEVANZ

Mit dem Kriterium der wissenschaftlichen Relevanz werden die vor allem die Bedeutung der erhobenen Ergebnisse für den Anwender und die Unabhängigkeit der Befragungsergebnisse von Evaluator und Versuchsleiter sowie die Ergebnisübereinstimmung von Evaluationen desselben Systems mit derselben Methode beschrieben. Bezogen auf Fragebogenverfahren werden mit dem Kriterium der wissenschaftlichen Relevanz diese zum Teil unter den Gesichtspunkten der Hauptgütekriterien bei der Testkonstruktion abgefragt. Eine allgemeine Beurteilung von Fragebogen-verfahren bezüglich der Hauptgütekriterien fällt zunächst schwer, da diese zum Teil stark variieren, bzw. je nach Fragebogen in unterschiedlichem Maße erfüllt sein können. Es lässt sich jedoch festhalten, dass Online-Fragebogenverfahren sich durch ein hohes Maß an Unabhängigkeit gegenüber dem Evaluator (räumliche sowie teilweise zeitliche Trennung von

Evaluator und Proband) auszeichnen und für den Anwender im Kontext der Benutzerzufriedenheit einen hohen praktischen Nutzen aufweisen (Aufdeckung von Usability-Problemen und anschließende Behebung). Fragebogenverfahren stellen (z.B. im Vergleich zum Interview) eine stark formalisierte Methode der Erhebung von Nutzerdaten dar und gewährleisten so bereits per se ein gewisses Ausmaß an Objektivität, welches durch eine automatisierte Auswertung noch verstärkt werden kann. Die Reliabilität eines (Online-) Fragebogens muss schließlich praktisch getestet werden, um gültige Aussagen treffen zu können[vgl. Sarodnick, Brau 2003:185ff.]. Nachfolgend sind die Kriterien der wissenschaftlichen Relevanz schematisch dargestellt:

Faktoren	Erklärung	Beurteilung von Fragebogenerfahren
Vorhersagekraft	Relevanz der Ergebnisse für den Anwender	Hoch, da konkrete Aussagen über Zufriedenheit der Nutzer und Gebrauchstauglichkeit der Website
Evaluator-Effekt	Unabhängigkeit der Ergebnisse von den Evaluatoren	Hoch, da räumliche und u.U. zeitliche Trennung von Evaluator und Befragtem
Objektivität	Unabhängigkeit der Ergebnisse vom Versuchsleiter	Hoch (vergleichbar Evaluator-Effekt)
Reliabilität	Ausmaß der Ergebnisübereinstimmung bei mehrmaliger Testdurchführung unter denselben Bedingungen	Abhängig von der Konstruktion des Fragebogens

**Tabelle 3: Wissenschaftliche Relevanz von Fragebogenverfahren in der Usability-Evaluation**

## 5.4 FAZIT

Es konnte gezeigt werden, dass Verfahren zur Datenerhebung mittels Online-Befragungen unter wissenschaftlichen und praktischen Gesichtspunkten ein relevantes Instrument zur Usability-Evaluation von Websites darstellen. Die beschriebenen Vorteile zeichnen die Online-Befragung sowohl unter qualitätstechnischen als auch unter aufwandstechnischen Kriterien als geeignetes Mittel aus, um mit begrenztem Aufwand aussagekräftige Daten erheben zu können. Allerdings sollte stets die eingeschränkte Repräsentativität der Ergebnisdaten von Online-Befragungen in Bezug auf eine größere Zielgruppe kritisch betrachtet werden. Für die vorliegende Arbeit, die sich mit der Konstruktion von Usability-Fragebögen jedoch nicht zum Ziel macht, repräsentative Daten für eine größere Stichprobe zu erheben, sondern Hinweise auf Probleme in der Gebrauchstauglichkeit von Websites aufzudecken, stellt diese Einschränkung bei einer ausreichend großen Stichprobe jedoch einen akzeptablen Umstand dar. Die Tatsache, dass Nutzertests in der Usability-Evaluation nicht als einzige Methode, sondern immer zusammen mit

expertenbasierten Methoden eingesetzt werden sollten, sollte gewährleisten, dass die begrenzte Repräsentativität der erhobenen Informationen durch die Daten einer zweiten Evaluationsmethode abgesichert werden kann.

Nachdem im vorangegangenen Abschnitt nach einem kurzen Überblick über Testtheorie und Gütekriterien die Relevanz von Fragebogenverfahren in der Usability-Evaluation dargelegt wurde, sollen im Anschluss bereits bestehende Fragebögen, die für diesen Zweck entwickelt wurden, vorgestellt werden. Dies geschieht mit dem Zweck, die eigenen, später zu entwickelnden Items bezüglich Skalenwahl und Itemformulierung in einen bereits bestehenden Kontext zu stellen und in diesem entsprechend zu diskutieren.

## 6 BESTEHENDE FRAGEBÖGEN ZUR ONLINE-NUTZERBEFRAGUNG

Um Nutzerdaten mit Fragebögen online zu erheben, stehen mehrere standardisierte und vielfach getestete Fragebögen zur Verfügung, die sich in Skalenwahl, Nutzungskontext und Zielgruppe zum Teil unterscheiden. Für die in Abschnitt 8 folgende Erstellung der Items sollen einige der bestehenden Fragebögen vorgestellt und näher diskutiert werden, um die Adaptierbarkeit von Skalenwahl und Itemformulierung auf den zu erstellenden Pool zu ermitteln. Die vorgestellten Fragebögen zur Erhebung von Software- und Website-Usability sind:

- QUIS (Questionnaire for User Interface Satisfaction)
- SUMI (Software Usability Measurement Inventory)
- Isometrics
- IsoNorm 9241
- WAMMI (Website Analysis and Measurement Inventory)
- AttrakDiff
- UEQ (User Experience Questionnaire)

### 6.1 QUIS (QUESTIONNAIRE FOR USER INTERFACE SATISFACTION)

Der Fragebogen zur Messung von Nutzerzufriedenheit in der Interaktion mit Benutzeroberflächen (Questionnaire for User Interaction Satisfaction) wurde bereits gegen Ende der 80er Jahre entwickelt und zielt speziell auf die Zufriedenheit eines Nutzers mit einer spezifischen Oberfläche ab. Der QUIS erhebt neben einer soziodemographischen Skala mit sechs Subskalen einen Messwert der Gesamtzufriedenheit sowie vier Messwerte zur spezifisch oberflächenorientierten Zufriedenheit (Bildschirmeigenschaften, Terminologie, Lernförderlichkeit, Systemressourcen). Die Skalen im QUIS sind jeweils neunstufig. Der QUIS liegt als Online-Version vor und wird vor allem in Lehre und Forschung (56%) sowie im kommerziellen Bereich (44%) eingesetzt [Norman, Harper 1993: 224f.].

Overall Reaction to the Software										
	1	2	3	4	5	6	7	8	9	
terrible										wonderful

Characters on the Computer Screen										
	1	2	3	4	5	6	7	8	9	
Hard to read										Easy to read

Abbildung 1: QUIS-Skala

[Chin et al. 1988:215]

## 6.2 SUMI (SOFTWARE USABILITY MEASUREMENT INVENTORY)

Das sogenannte *Software Usability Measurement Inventory* wurde Mitte der 90er Jahre veröffentlicht und versucht, subjektive Nutzereindrücke und Gefühle bei der Interaktion mit Software zu erfassen. Der Fragebogen setzt sich aus 50 Fragebogenitems zusammen, die in Aussagenform aufgebaut sind und denen der Nutzer entweder zustimmen, nicht zustimmen, oder sich unentschieden zeigen kann. Die Standardisierungsdatenbank besteht aus über 1000 beantworteten Fragebögen, der SUMI gehört somit zu den etabliertesten und bekanntesten Fragebögen, der zur Evaluation von Software besteht, ist jedoch kostenpflichtig. SUMI verwendet keine Subskalen, und Sarodnick und Brau merken kritisch an, dass der Fragebogen keinen Bezug zur ISO-Norm 9241 aufweist [s. Sarodnick, Brau 2003:176].

		Disagree	Undecided	Agree
		↓	↓	↓
18	There is never enough information on the screen when it's needed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19	I feel in command of this software when I am using it.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	I prefer to stick to the facilities that I know best.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 2: SUMI – Skala, UK-Sample

[<http://sumi.ucc.ie/uksample.pdf>, Zugriff 22.03.2010, 07:51 ME0Z]

### 6.3 ISOMETRICS

Ebenso wie der SUMI Mitte der 90er Jahre entstanden, stellt der IsoMetrics ein in Anlehnung an die internationale ISO-Norm 9241 entwickeltes Instrument zur Evaluation von Software dar. Im Gegensatz zum SUMI mit drei Skalenabstufungen enthält der IsoMetrics eine fünfstufige Likert-Skala (Zustimmung-keine Zustimmung) sowie die Möglichkeit für den Nutzer, sich nicht zu einem Item zu äußern. Der IsoMetrics liegt in Kurz- (reine Items) und in Langform (Möglichkeit für den Evaluierenden, die individuelle Relevanz von Items zu gewichten) vor. IsoMetrics besteht aus 90 Items, die nach Auswertung eine detaillierte Auflistung von Usability-Problemen ermöglichen. Der IsoMetrics wurde vielfach getestet und weist ebenso wie der SUMI einen hohen Grad an Reliabilität und Validität auf, er liegt sowohl als Online- als auch als Papierform vor [s. Sarodnick, Brau 2003:178f.].

		stimmt nicht	stimmt wenig	stimmt mittelmäßig	stimmt ziemlich	stimmt sehr	keine Angabe
A.1	Die Software zwingt mich, überflüssige Arbeitsschritte durchzuführen						

	nicht wichtig	wenig wichtig	mittelmäßig wichtig	ziemlich wichtig	sehr wichtig	keine Angabe
Wie wichtig ist Ihnen der Aspekt für Ihren Gesamteindruck von dieser Software?						

Abbildung 3: IsoMetrics-Skala

[s. Gediga et al. 1999:154]

### 6.4 ISONORM 9241

Der IsoNorm Fragebogen der Autoren Prümper und Anft zur Erhebung der Gebrauchstauglichkeit von Software am Arbeitsplatz beinhaltet in Anlehnung an die ISO-Norm 9241 sieben Subskalen, die sich thematisch in die Anforderungen zur Dialoggestaltung von Systemen gliedern lassen. Jeweils fünf Items pro Subskala überprüfen die Gebrauchstauglichkeit des zu evaluierenden Systems, ein Frageteil zur Erhebung soziodemographischer Daten ist im Anhang angeschlossen.

IsoNorm beinhaltet eine siebenstufige Skala, auf der der Evaluator das zu evaluierende System in einem semantischen Differential von negativ bis positiv bewerten kann, Hinweise auf Schwachstellen des Systems können so schnell und effizient aufgezeigt werden. IsoNorm ist als Papierversion und als Online-Version erhältlich [s. Sarodnick, Brau 2003:177f.].

Die Software...	---	--	-	-/+	+	++	+++	Die Software
...ist kompliziert zu bedienen								...ist unkompliziert zu bedienen

Abbildung 4: IsoNorm 9241

[s. Prümper, Anft 1993:3]

## 6.5 WAMMI (WEBSITE ANALYSIS AND MEASUREMENT INVENTORY)

Das Website Analysis and Measurement Inventory (WAMMI) ist im Gegensatz zu den zuvor beschriebenen Methoden speziell auf die Evaluation von Websites zugeschnitten und wurde Ende der 90er Jahre aus dem SUMI entwickelt. Einzelne Items wurden direkt übernommen und auf Websites angepasst, was sich geringfügig auf die innere Konsistenz des Fragebogens und der Subskalen auswirkt, wie Thielsch anmerkt. Mit Hilfe von fünf Subskalen à vier Items erfasst der WAMMI ebenso wie der SUMI die Einstellung und Meinungen der Nutzer durch deren Zustimmung oder Ablehnung zu insgesamt 20 Aussagen. Der WAMMI wird wie der SUMI kommerziell (kostenpflichtig) und international vertrieben, Angaben über Validität und Reliabilität sind nur begrenzt zugänglich [s. Thielsch 2008:19].

						Strongly agree ↓	Strongly disagree ↓
This Website has much that is of interest for me							

Abbildung 5: WAMMI – Skala

[<http://www.wammi.com/samples/index.html>, Zugriff 31. März 2010]

## 6.6 ATTRAKDIFF

Im Gegensatz zur den bisher beschriebenen standardisierte Fragebögen, in denen zumeist die Gebrauchstauglichkeit (Effektivität, Effizienz) eines Systems evaluiert wird, setzt sich der 2000 entwickelte Fragebogen AttrakDiff von Hassenzahl, Burmester und Koller mit dem eher

subjektiven Konzept der Nutzerzufriedenheit und der sogenannten „joy of use“ bei der Benutzung eines Systems auseinander. Hassenzahl et al. unterteilen hierbei in pragmatische (Gebrauchstauglichkeit) und hedonische (Nutzungserlebnis) Qualität. Die hedonische Qualität eines Systems setzt sich hierbei aus Stimulation (Forderung des Nutzers durch das Produkt oder die Anwendung, Lerneffekte) und Identität (Vermittlung eines Selbstwerts im Gebrauch des Produktes/der Anwendung) zusammen. Zur Erhebung der Nutzerzufriedenheit und der Gebrauchstauglichkeit des Systems werden im AttrakDiff 28 bipolare Items in semantischen Differentialen mittels einer siebenstufigen Skala verwendet. AttrakDiff liegt wie der IsoMetrics- und der IsoNorm-Fragebogen ebenfalls auch als Online-Version vor [Hassenzahl et al. 2008:79f.].

Bitte geben Sie mit Hilfe der folgenden Wortpaare Ihren Eindruck des Produkts wieder. Bitte kreuzen Sie nur ein Kästchen an!								
angenehm	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unangenehm
Originell	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	konventionell

Abbildung 6: Semantisches Differential des AttrakDiff

[s. Sarodnick, Brau 2003:180]

## 6.7 UEQ (USER EXPERIENCE QUESTIONNAIRE)

Der User Experience Questionnaire wurde 2006 im Rahmen der SAP AG entwickelt und stellt ein Instrument dar, mit dem verschiedene Aspekte der Qualität von Softwareprodukten erhoben werden können. Diese werden in 26 Items auf sechs Subskalen (Attraktivität, Durchschaubarkeit, Effizienz, Vorhersagbarkeit, Stimulation, Originalität) abgebildet. Die Einstellung der Benutzer zu der zu evaluierenden Software wird mit Hilfe eines siebenstufigen semantischen Differentials erhoben. Bezüglich Validierung und Reliabilität liegen aufgrund der geringen Größe der Testgruppen (13-16 Teilnehmer) bisher nur begrenzt aussagekräftige Ergebnisse aus kleineren Usability-Studien vor.

Einschätzung des Produktes								
	1	2	3	4	5	6	7	
attraktiv	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unattraktiv

Abbildung 7: User Experience Questionnaire

[vgl. Laugwitz et al. 2006:125ff.]



## 6.8 ZUSAMMENFASSUNG

Zur Bewertung von Systemen und Anwendungen existieren einige, bereits standardisierte, sowie auf Reliabilität und Validität getestete Fragebögen. Da zumeist für den Einsatz in der Interaktion mit Software entwickelt, sind diese bezüglich Ausführlichkeit und Länge recht umfangreich. In neuer entwickelten Instrumenten wie dem WAMMI, die speziell für den Einsatz in der Evaluation von Websites entwickelt wurden, lässt sich der Trend zu einer Skalenverkürzung und deutlichen Verringerung der Itemanzahl erkennen. Wie bereits im vorhergehenden Abschnitt erläutert, ist dies ein notwendiges Vorgehen, da die Nonresponse-Quote der Teilnehmer mit zunehmender Anzahl an Fragen deutlich ansteigt [vgl. Kaczmirek 2008: 40f.].

Von den beschriebenen sieben Fragebogenverfahren weisen vier Skalen in Form eines semantischen Differentials auf (QUIS, IsoNorm, AttrakDiff, UEQ). Die restlichen drei der beschriebenen Bogen sind in Form von Likert-Skalen (Zustimmung zu vorgegebenen Aussagen) konzeptioniert. Die Items liegen in überwiegender Form in einer Positivformulierung vor.

Bezüglich der Entwicklung des geplanten Fragebogengenerators ist festzuhalten, dass die Länge des vom Anwender zu erstellenden Fragebogens sowie die Auswahlskala deutlich begrenzt sein sollten, um Scrollen zu vermeiden und Übersichtlichkeit zu gewährleisten. In Bezug auf die Skalenwahl lassen sich keine eindeutigen Tendenzen zum semantischen Differential oder zur Likert-Skala ableiten. Überlegungen zu diesem Aspekt sollen in Abschnitt 8.1 bei der Itemgenerierung weiter beschrieben werden.

## 6.9 FAZIT

Im vorangegangenen Kapitel konnte ein ausführlicher Überblick über verschiedene bestehende Fragebögen zur Usability-Evaluation geliefert werden. Zur Entwicklung des geplanten Fragebogengenerators lieferte die Analyse von Itemformulierung, Skalenwahl und Fragebogenlänge wichtige Hinweise, die in die Entwicklung der späteren Items mit einfließen können.

Im Folgenden soll eine Kategorisierung unterschiedlicher Website-Typen vorgenommen werden, die bei der späteren Erstellung des Fragebogengenerators dazu dienen soll, mithilfe websitespezifischer Items die Adaptierbarkeit des Tools zu gewährleisten. Dazu sollen zunächst bestehende Klassifikationsansätze beschrieben, eine eigene Klassifikation entwickelt, sowie im Anschluss unterschiedliche Anforderungen an die Usability der verschiedenen Kontexte analysiert werden.

## 7 KLASSIFIKATION VON WEBSITES UND UNTERSCHIEDLICHE ANWENDUNGSKONTEXTE

Das für das BUI-Projekt zu entwickelnde Fragebogentool soll die Evaluation der Usability von Webauftritten und Projekten in unterschiedlichen Anwendungskontexten ermöglichen. Um die spätere Adaptierbarkeit des Tools gewährleisten zu können, müssen im Vorfeld mögliche Anwendungskontexte, deren Gegenstand abgedeckt werden soll, definiert und beschrieben werden. Im Folgenden sollen daher die Klassifikation und Beschreibung unterschiedlicher Anwendungskontexte erfolgen. Typische Usability-Probleme, die in den ermittelten unterschiedlichen Website-Kategorien auftreten können, sollen aufgezeigt und analysiert werden.

### Websites im World Wide Web

Um eine wissenschaftliche Einordnung von Websites und Website-Typen vornehmen zu können, soll zunächst definiert werden, was unter den Begriffen Website und World Wide Web zu verstehen ist. Unter World Wide Web kann ein Hypertextsystem, das online zugänglich ist, verstanden werden. Der üblicherweise synonym verwandte Begriff Internet bezeichnet hingegen

*„ ... nur den Verbund verschiedener Computernetzwerke...“*

[Thielsch 2008:3]

auf denen diese Hypertextsysteme abrufbar sind. Zu unterscheiden ist hierbei zwischen dem Begriff Webseite<sup>15</sup> und dem Begriff der Website<sup>16</sup>. Eine Website enthält üblicherweise eine oder mehrere Webseiten, die wiederum hierarchisch angeordnet sein können.

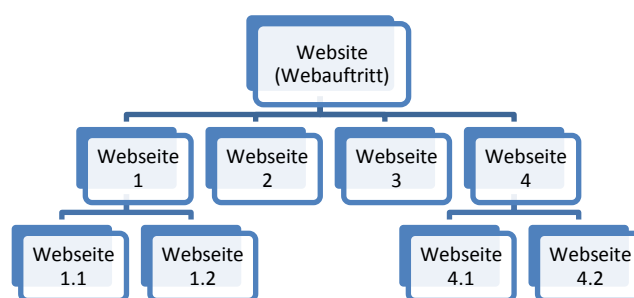


Abbildung 8: Struktur und Aufbau von Websites nach Thielsch 2008

Da das World Wide Web einen Umfang von geschätzt 22 Milliarden Webseiten aufweist und sich somit durch eine hohe Anzahl an unterschiedlichen Zielsetzungen und Funktionalitäten

<sup>15</sup> einzelne, mit HTML erstellte Seite innerhalb des WWW

<sup>16</sup> auch Webauftritt oder Internetpräsenz, Gesamtheit aller Webseiten eines Akteurs innerhalb eines inhaltlich eingegrenzten Bereichs

auszeichnet [vgl. a.a.O.:85], ist für die Entwicklung eines Tools mit dem Ziel der Erhebung der Benutzerzufriedenheit zunächst eine thematisch eingrenzende Klassifikation der unterschiedlichen Typen von Websites nötig.

## 7.1 KLASSIFIKATION UNTERSCHIEDLICHER WEBSITE-TYPEN

Websites werden in der Literatur unter verschiedenen Gesichtspunkten kategorisiert, hierbei kann beispielweise anhand der Kriterien Zugriffshöhe, Nutzerzahlen, Funktionalitäten oder Anwendungskontext vorgegangen werden [vgl. a.a.O.:85f.]. Im Folgenden sollen Websites bezogen auf das Klassifikationsmerkmal „Anwendungskontext“ charakterisiert werden. Hierzu sollen zunächst ausgewählte bestehende Klassifikationen aus der Literatur vorgestellt werden, woraufhin im Anschluss ein eigenes Kategorisierungsschema unter Berücksichtigung bestehender Ansätze erstellt werden soll.

### 7.1.1 METHODISCHES VORGEHEN

Die Anzahl von Klassifikationsansätzen ist umfassend, in der Literatur jedoch häufig auf den Bereich E-Business/E-Commerce (zur Begriffsklärung s. Abschnitt 7.4.2) fokussiert. Allgemeinere Klassifikationen, die sich um den Versuch bemühen, den Gesamtbereich aller Websites zu schematisieren, sind seltener, was möglicherweise auf die Schwierigkeit der Aufgabe in Bezug auf den Umfang und die stetig variierende Natur des Mediums World Wide Web zurückzuführen sein kann. Dennoch gibt es eine Vielzahl an Klassifikationsansätzen, von denen zum einen aus Gründen des Umfangs der Arbeit, zum anderen aus Ähnlichkeitsgründen der Klassifikationsschemata (v.a. im Bereich E-Commerce) nur einige ausgewählte Kategorisierungsansätze vorgestellt werden sollen. Hierbei wurde vor allem darauf geachtet, Schemata aus einem umfassenden Zeitraum (90er Jahre bis zum aktuellen Zeitpunkt) auszuwählen. Diese ordnen Websites unter verschiedenen Gesichtspunkten ein, um die Vielzahl an möglichen Ansätzen und Betrachtungsweisen effektiv und anschaulich darstellen zu können. Die Auswahl der Kategorisierungen erhebt demnach keinen Anspruch auf Vollständigkeit. Da sie jedoch nur als Anhaltspunkt für die Entwicklung eines eigenen Klassifikationsansatzes dienen soll, kann dies als gerechtfertigt betrachtet werden.

### 7.1.2 BESTEHENDE KLASSIFIKATIONSANSÄTZE

Das „Project 2000: Research Program on Marketing in Computer-Mediated Environments“ der Vanderbilt University (Nashville, Tennessee) setzt sich seit Mitte der 90er Jahre zum Ziel, die Bedeutung des Mediums World Wide Web vor kommerziellem Hintergrund zu erforschen und zu analysieren. Hoffmann et al. (1995) klassifizieren in ihrer Arbeit, die sich mit der Bedeutung sowie der Entwicklung des Web im diesem Kontext auseinandersetzt, Websites im Bereich E-Business/E-Commerce gemäß deren Zielsetzung, und kommen zu folgenden sechs Kategorien:

- Online-Storefront (direkter Vertrieb via Internet)
- Internet-Präsenz (Werbung, Bilder, Information)
- Inhaltsbasierte Seiten (Bereitstellung von kostenpflichtigem/nichtkostenpflichtigem Inhalt)
- Mall (Zusammenschluss mehrerer Online-Storefronts)
- Incentive-Seiten (Werbung für zugehörige kommerzielle Seiten)
- Suchmaschinen

[Hoffmann et al. 1995:12ff.]

Hong und Kim (2004) gehen in ihrer Arbeit zur Entwicklung eines allgemeinen Rahmenwerks zur systematischen Evaluation von Websites nach ihrem jeweiligen Typ mit einer dichotomen Klassifikation (Nutzerziel und Nutzeraktivität) vor und spezifizieren insgesamt 16 Kategorien von Website-Typen im World Wide Web [Hong, Kim 2004:344f.]. Zwischen diesen herrschen jedoch z.T. hierarchische Beziehungen (General Shopping Mall → Speciality Shopping Mall) oder Überschneidungen (Organization Homepage / Economy & Industry):

No	Code	Category
1	VA1	General shopping mall
2	VA2	Specialty shopping mall
3	VA3	Auction
4	VA4	Reservation service
5	VA7	Health and medical
6	VA8	Computer and Internet
7	VA9	Economy and industry
8	VA10	Women and children
9	VA11	Games
10	VA12	Portals
11	VA13	Web hosting
12	VA14	Community
13	VA15	Personal homepages
14	VA16	Organization homepage
15	VA17	Online education
16	VA18	Web casting

Abbildung 9: Kategorisierung des Web nach Hong, Kim (2004)

De Marsico und Levaldi (2004) nehmen in ihrer Arbeit zur Evaluation von Websites mit dem Schwerpunkt auf der Erfüllung von Nutzererwartungen eine sehr viel allgemeinere Einteilung von Websites vor. Sie kategorisieren diese unter Einbeziehung des sogenannten rhetorischen Dreiecks

und kommen auf eine Taxonomie mit insgesamt neun eher abstrakten Kategorien im dreidimensionalen Raum [Levialdi S. DeMarsico M. 2004:399].

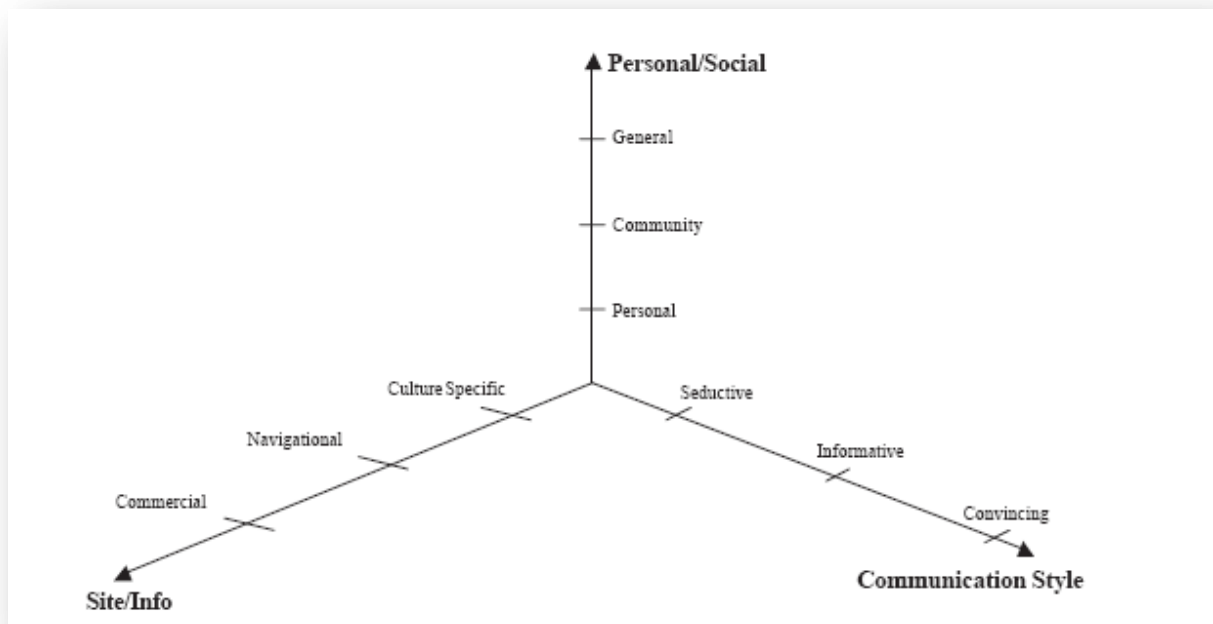


Abbildung 10: Kategorisierung des Web nach Levialdi, DeMarsico (2004)

Zviran et al. (2005) beschäftigen sich in ihrer Arbeit zur Nutzerzufriedenheit auf Websites mit kommerziellem Hintergrund in der Zeitschrift *Information and Management* ebenfalls mit dem Feld E-Business/E-Commerce und übernehmen die IBM-Webklassifikation, die Websites anhand ihrer Besucherzahlen unter den weiter gehenden Kriterien Transaktionszahlen, Komplexität der Website, Suchzahlen, Stabilität der Information sowie Sicherheitsaspekte in die folgenden Kategorien einordnet:

- Publish/Subscribe (Bereitstellung von Informationen für Nutzer, Suchmaschinen & Informationsseiten)
- Online shopping
- Customer self-service (Bereitstellung von Online-Funktionalitäten für Nutzer, z.B. Online-Banking, Reiseagenturen)
- Trading (Nutzer können sowohl kaufen als auch verkaufen, bspw. Online-Auktionshäuser)
- B2B (Business to Business-Handelsbeziehungen)

[Zviran et al. 2005:158ff.]

Thielsch (2008) identifiziert in seiner Arbeit insgesamt zehn Kategorien, basierend auf einer Klassifikation aus „inhaltlich-funktionalen Aspekten“, und unterteilt Websites anhand von Inhalt (Information, Produkte, etc.) und Funktionalität (Kommunikation, Handel, etc.).

- E-Commerce
- Entertainment
- E-Learning
- E-Recruiting/E-Assessment
- Information
- Kommunikation/Online-Netzwerke
- Portalfunktion
- Selbstdarstellung
- Social Software
- Suchmaschinen

[Thielsch 2008:86]

### 7.1.3 ZUSAMMENFASSUNG

Im vorangegangenen Teil wurden insgesamt fünf unterschiedliche Klassifikationsansätze vorgestellt, die veranschaulichen sollen, dass die Kategorisierung von Websites einen stark subjektiven Charakter aufweist, und vor allem auf der Schwerpunktsetzung des Autors beruht. Eine allgemeine Einteilung in Website-Typen kann – wie bei Hong und Kim (2004), Thielsch (2008) oder DeMarsico und Levialdi (2004) gezeigt – in unterschiedlichem Detaillierungsgrad und anhand verschiedener Kriterien vorgenommen werden. Einzelne Bereiche, wie am Beispiel von Hoffman et al. (1995) und Zviran et al. (2005) verdeutlicht, sind ebenfalls weiter und in beliebiger Spezifiziertheit unterteilbar. Nachfolgend eine zusammenfassende Aufstellung der verschiedenen Klassifikationsansätze:

Autor	Jahr	Thematik	Anzahl Kategorien
Hoffmann et al.	1995	E-Business/E-Commerce	6
Hong & Kim	2004	Allgemein	16
De Marsico & Levialdi	2004	Erfüllung von Nutzererwartungen	9
Zviran et al.	2005	E-Business/E-Commerce	5
Thielsch	2008	Allgemeine Klassifikation	10

Tabelle 4: Klassifikationsansätze nach unterschiedlichen Autoren

In Anlehnung an die bereits bestehenden Klassifikationsansätze soll nun im folgenden Abschnitt eine eigene Kategorisierung von Websites erstellt werden.

## 7.2 ENTWICKLUNG EINER EIGENEN KLASSIFIKATION ZUM EINSATZ IM FRAGEBOGENGENERATOR

Wie bereits am Anfang des Kapitels festgehalten, erschwert die zunehmende Komplexität des World Wide Web sowie seine variierende Natur eine eindeutige und unstrittige Kategorisierung. Trotz erkennbarer thematischer Grundstrukturen (Information, Kommunikation, Handel, etc.) verlangt die umfassende Aufgabe einer Einteilung des World Wide Webs stets nach einem pragmatischen, der jeweiligen Erkenntnisorientierung angepassten Kategorisierungsschema, mit dem im betreffenden Anwendungskontext vorgegangen werden kann. Entsprechend muss für das im Rahmen der Arbeit zu entwickelnde Fragebogentool eine eigene Klassifikation entworfen werden. Da die bestehenden Klassifikationen von Hoffmann et al. (1995) sowie Zviran et al. (2005) lediglich den Bereich E-Business/E-Commerce weiter unterteilen, und die Klassifikation von DeMarsico und Levialdi (2004) für den konkreten Gebrauch zu abstrakt erscheint, bleiben die Klassifikationen von Hong und Kim (2004) sowie Thielsch (2008), die für die Zielsetzung der Arbeit als am ehesten thematisch geeignet erscheinen, zu analysieren. Im Anschluss sollen aus dem Vergleich der beiden Klassifikationsansätze eigene Kategorien entwickelt werden, die im zu erstellenden Fragebogentool Verwendung finden sollen. Nachfolgend eine Gegenüberstellung der beiden Klassifikationen:

Hong & Kim 2004	Thielsch 2008
General Shopping Mall	E-Commerce
Specialty Shopping Mall	E-Commerce
Auction	E-Commerce
Reservation Service	k.E. <sup>17</sup>
Health and Medical	k.E.
Computer and Internet	k.E.
Economy and Industry	k.E.
Woman and Children	k.E.
Games	Entertainment
Portals	Portalfunktion
Web Hosting	k. E.
Community	Kommunikation/Online-Netzwerke/Social Software
Personal Homepage	Selbstdarstellung, Information
Organization Homepage	Selbstdarstellung, Information
Online Education	E-Learning
Web Casting	Entertainment
k.E.	E-Recruiting/E-Assessment
k.E.	Suchmaschinen

Tabelle 5: Klassifikationen von Websites nach Thielsch (2008) /Hong&Kim (2004)

<sup>17</sup> Keine Entsprechung

An dieser Stelle wurden die Kategorien von Hong und Kim als Vergleichspunkt gewählt und bei Thielsch nach jeweiligen Entsprechungen gesucht. Ein Vorgehen in umgekehrter Reihenfolge (s. Appendix A) kann ebenso vorgenommen werden und führt zu denselben Ergebnissen.

Im Vergleich der beiden Klassifikationen finden sich häufig Entsprechungen in den allgemeinen Kategorien, so konnten Übereinstimmungen in den Bereichen

- E-Commerce – Shopping Mall (General/Specialty), Auction
- Entertainment – Games, Webcasting
- Portalfunktionen – Portals
- Social Software, Online-Netzwerke – Communities
- Selbstdarstellung, Information – Homepage (Organizational/Personal)
- E-Learning – Online-Education

gefunden werden. Da die Klassifikation von Hong und Kim mit 16 Kategorien zum einen als zu umfangreich erscheint sowie einen starken Spezialisierungsgrad aufweist, wurden aus ihr die Kategorien, die keine inhaltliche Entsprechung mit der Kategorisierung von Thielsch besitzen (Health & Medical, Economy & Industry, Web Hosting, Woman & Children, Reservation Service, Computer & Internet) entfernt. Zum einen ist aus Gründen des Umfangs der Arbeit die Anzahl der möglichen Kategorien zu begrenzen, zum anderen können bei der Arbeit mit dem späteren Fragebogengenerators die entsprechenden Kategorien hierarchisch als Subkategorien in anderen Bereichen eingeordnet werden. So könnten Health & Medical sowie Woman & Children den Oberkategorien Informative Seiten und/oder Portalen zugeordnet werden, Economy & Industry kann entweder der Kategorie Selbstdarstellung (Unternehmenshomepage) oder E-Commerce zugeordnet werden, ebenso wie Reservation Services und Web Hosting. Der Bereich Computer und Internet ist so vage gefasst, dass er – je nach Content der Website – entweder der Kategorie Portale,

E-Commerce oder Selbstdarstellung (Unternehmenshomepage) zuzuordnen ist.

Konjunkte Klassen wurden daher aus Gründen der Operationalisierbarkeit zusammengefasst. So wurden beispielsweise die Kategorien Communities, Social Software, Netzwerke und Kommunikation zusammen mit dem Bereich E-Recruiting zur Kategorie Portale zusammengelegt, da sie gemeinsame Grundfunktionen (Anmeldung der User mit Nutzerprofil, Kommunikationsmöglichkeiten über Foren und private Nachrichten) aufweisen. Shopping Mall (General und Specialty) wurden zusammen mit Auctions zur Kategorie E-Commerce gefasst, Games und Webcasting zur Kategorie Entertainment“. Private sowie kommerzielle Homepages fallen unter den Bereich „Selbstdarstellung“ (Repräsentation).

Zusammengefasst ergeben sich als Klassifikationsrahmen für das zu erstellende Fragebogentool die sechs Kategorien

- E-Commerce (Online-Shops, Auktionen)
- Entertainment (Games, Webradio, WebTV)



- Portale (Communities, Social Software, Online-Netzwerke, Kommunikation, E-Recruiting)
- Information (informative Seiten, z.B. Online-Auftritte von Zeitungen/Zeitschriften, Online-Enzyklopädien)
- Präsentation (von z.B. Unternehmen/Organisationen)
- E-Learning

### 7.3 ZUSAMMENFASSUNG

In Anlehnung an bereits bestehende Ansätze zur Klassifizierung konnte im vorangegangenen Teil ein eigenes Kategorisierungsschema für Websites unterschiedlicher Anwendungskontexte entwickelt werden. Ergebnis des Vorgehens waren die sechs Kategorien E-Commerce, Entertainment, Portale, Information, Präsentation und E-Learning. Typische Beispiele für Webseiten seien im Folgenden schematisch dargestellt:

Kategorie	Beispiel
E-Commerce	amazon.de, ebay.de
Entertainment	youtube.com, myvideo.de, sevengames.de
Portale	facebook.com, monster.de
Information	yahoo.de, sueddeutsche.de, bhf-bank.com
Präsentation	mcdonalds.de, uni-hildesheim.de
E-Learning	http://video.tu-clausthal.de, www.bildung.at

Tabelle 6: Beispielwebsites für die ermittelten Kategorien

Unter Berücksichtigung der Zielsetzung und des Umfangs der Arbeit wurden die einzelnen Kategorien des Schemas zum Teil deutlich gröber gesetzt, als bei bereits bestehenden Klassifikationen. Einige Kategorien wurden komplett weggelassen, teils aus Gründen der Spezifiziertheit (z.B. Health and Medical), teils aufgrund der Zusammenfassung zu größeren Clustern (vgl. Social Software und Online-Netzwerke zu Portalen), teils aufgrund der Zielgruppe<sup>18</sup> der Plattform [www.usability-toolkit.de](http://www.usability-toolkit.de) (bspw. Wegfall der Klasse Suchmaschine). Auch die denkbaren, möglicherweise extra zu fassenden Klassen E-Banking und E-Government werden aufgrund der Zielgruppenspezifikation hier nicht explizit als Klassen mit aufgeführt, wären in anderem Kontext jedoch zu nennen. Da für jede Kategorie eine eigene Analyse auf konkrete Usability-Probleme und Anforderungen an die Gebrauchstauglichkeit sowie der Entwurf von konkreten, kategoriespezifischen Fragebogenitems vorgenommen werden muss, war vor allem die Zusammenfassung von mehreren Unterkategorien zu größeren Clustern aus Gründen des Umfangs der Arbeit sinnvoll.

Festzuhalten sei an dieser Stelle, dass die genannten Kategorien selbstverständlich nicht immer überschneidungsfrei sind, da beispielsweise Informationsseiten wie Online-Zeitungen

<sup>18</sup> Kleine und mittelständische Unternehmen

(sueddeutsche.de, faz.net) meist ebenfalls einen Online-Shop bieten, und selbstdarstellende Seiten (Unternehmenswebsites, Websites von Hochschulen) immer auch Informationsgehalt für den Nutzer bieten. Der Trend zu immer multidimensionalen Seiten, die dem Nutzer umfassende Funktionskataloge bieten (Information, Shop, Kommunikation mit anderen Nutzern, Profilerstellung), erschwert eine klare und unstrittige Kategorisierung. Die Websites sollen an dieser Stelle entsprechend ihrer vorherrschenden inhaltlichen Ausrichtung den einzelnen Kategorien zugeordnet werden. Die Abgrenzung insbesondere von informativen zu repräsentativen Seiten wurde hier über die Bereitstellung von zeitvarianten Informationen (regelmäßige Aktualisierung des Contents in bspw. Online-Zeitungen) versus (größtenteils) zeitinvariantem Informationsgehalt von beispielsweise Unternehmenswebsites vorgenommen.

Für die genannten Kategorien sollen nun im Folgenden spezifische Usability-Anforderungen ermittelt werden, anhand derer die Websites beurteilt werden können, und die die Grundlage für die im Anschluss zu erstellenden Fragebogenitems bilden sollen.

#### 7.4 SPEZIFISCHE USABILITY-ANFORDERUNGEN IN DEN UNTERSCHIEDLICHEN KATEGORIEN

Um Websites im Allgemeinen sowie in den genannten Kategorien im Speziellen beurteilen zu können, müssen allgemeingültige Messgrößen wie in Abschnitt 2.1 beschrieben, verwendet werden. Um Usability (Effizienz, Effektivität, Zufriedenheit) im Webkontext beurteilen zu können, werden in der Literatur vielfach die Grundsätze der Dialoggestaltung nach ISO-9241/10 herangezogen.

Diese lauten wie folgt:

- Aufgabenangemessenheit (Unterstützung für den Benutzer, seine Aufgaben effektiv und effizient zu erledigen)
- Selbstbeschreibungsfähigkeit (Rückmeldung des Systems bei Dialogschritten, Systemhilfe auf Anfrage)
- Steuerbarkeit (Möglichkeit des Nutzers, den Dialogablauf zu starten und in Geschwindigkeit und Richtung zu beeinflussen)
- Erwartungskonformität (Konsistenz, Entsprechung allgemeiner Normen, den Merkmalen der Nutzer angepasst)
- Fehlertoleranz (Fehlerhafte Eingaben können mit minimalem Korrekturaufwand berichtigt werden)
- Individualisierbarkeit (Anpassung des Dialogs an die entsprechende Aufgabe und/oder an die Bedürfnisse des Benutzers)
- Lernförderlichkeit (Das System unterstützt den Nutzer beim Erlernen des Systems und seiner Funktionen)

Fragebögen zur Usability-Evaluation, die die allgemeine Usability eines Produktes oder einer Website anhand von Kriterien, die aus den oben aufgeführten Grundsätzen ermittelt wurden, erheben wurden bereits in Abschnitt 6 beschrieben. Im Gegensatz zu diesen allgemeinen (generischen) Usability-Kriterien müssen in der vorliegenden Arbeit jedoch innerhalb der ermittelten Kategorien spezifische Anforderungen an die Gebrauchstauglichkeit für die einzelnen Klassen von Websites ermittelt werden, die zum Teil in die Anforderungen nach ISO 9241 einzuordnen sind, teilweise jedoch über die allgemeinen Grundätze der Dialoggestaltung hinausgehen.

---

#### 7.4.1 METHODISCHES VORGEHEN

Um die Adaptierbarkeit des zu erstellenden Fragebogentools zu gewährleisten, werden in der vorliegenden Arbeit categoriespezifische Usability-Kriterien aus der Literatur ermittelt. Diese variieren für jede Kategorie. So sind beispielsweise Anforderungen an Online-Shops ([www.amazon.de](http://www.amazon.de)) anderer Natur als Anforderungen an Seiten, die den Nutzer unterhalten ([www.youtube.com](http://www.youtube.com)) oder informieren sollen ([www.sueddeutsche.de](http://www.sueddeutsche.de)). Die ermittelten Usability-Kriterien dienen im Anschluss als Grundlage für die Erstellung der categoriespezifischen Fragebogenitems. Es wurde an dieser Stelle versucht, die spezifischen Usability-Kriterien möglichst umfassend und vollständig zu ermitteln. Einen Anspruch auf endgültige Vollständigkeit zu erheben, ist jedoch, bedingt vor allem durch den kontinuierlichen Entwicklungsprozess, dem das World Wide Web und die Forschung im Bereich Usability/User Experience unterliegen, nur schwer möglich.

---

#### 7.4.2 E-COMMERCE

Die Begriffe E-Commerce und E-Business umfassen die Verlagerung von Geschäftsprozessen (Transaktionen, Handel, Dienstleistung) in den Online-Kontext. E-Commerce bezeichnet hierbei den Geschäftsprozess zwischen Unternehmen und Kunde, E-Business denjenigen zwischen Unternehmen [s. Dohmann 2002:1ff.]. In der vorliegenden Arbeit liegt der Schwerpunkt daher entsprechend auf dem Bereich E-Commerce. Online-Shops wie [www.amazon.de](http://www.amazon.de) oder die Auktionsplattform [ebay.de](http://www.ebay.de) konnten in den letzten Jahren - bedingt durch die sprunghafte Verbreitung des Internet und die steigende Versorgung der Bevölkerung mit Breitbandanschlüssen - einen beispielhaften Zuwachs an Neukunden verzeichnen. 97% der deutschen Internetnutzer informieren sich über Produkte und Dienstleistungen im Internet, 66% der Bevölkerung haben bereits online eingekauft, wobei der proportionale Anteil der sogenannten „Intensivkäufer“ (10 und mehr Produkte im vorangegangenen Monat) ansteigend ist [s. SCHNELLER 2009:27]. Beliebte Produkte sind vor allem Bücher, Reisen und Hotelbuchungen sowie Kleidung und Computerhardware.

Um eine E-Commerce-Plattform erfolgreich gestalten zu können, sollte der Kaufprozess des Kunden möglichst realitätsgetreu abgebildet werden, wobei Bargas (2001) vor allem die Relevanz der Schritte des Kaufprozesses *Bedürfnisidentifikation* (welchen Wunsch hat der Kunde), *Produktvermittlung* (welches Produkt befriedigt die Anforderungen des Kunden), *Item Searching* (Suche nach Produkten, die den Anforderungen des Kunden gerecht werden) und *Kauf* (des gewünschten Produktes) hervorhebt. Produkte können entweder über *Existence Search* (der Kunde sucht nach Produkten, die in etwa seinen Anforderungen entsprechen) oder über *Known Item Search* (der Kunde sucht ein bestimmtes Produkt im Online-Shop) gefunden werden [vgl. Bargas-Avila 2001:23f.]. Die genannten Prozessschritte stellen unterschiedliche Anforderungen an die Usability der Seite: Sofern sich der Besucher des Online-Shops unschlüssig oder unsicher ist, wonach ein Bedürfnis besteht, kann vor allem die Übersichtlichkeit von Startseite und Hauptnavigation von Vorteil sein (Präsentation beliebter Produkte auf der Startseite bzw. geordnete Darstellung der angebotenen Produktkategorien in der Hauptnavigation [vgl. BVDW 2008(a):10f., Sauer, Schmeißer 2005:2]. Vor allem die Navigationsleiste auf der linken Seite sollte übersichtlich und klar gestaltet werden, da auf ihr die häufigsten Nutzerzugriffe zu verzeichnen sind. Eine begrenzte Anzahl von Kategorien trägt zur Orientierung des Nutzers bei [vgl. Puscher 2009:128, BVDW 2008(a):14]. Da viele Online-Shops auch horizontale Navigationsleisten verwenden, ist dies in der Itemformulierung unter Umständen auf eine sogenannte Hauptnavigation zu erweitern.

*Known Item Search* erfordert eine effiziente Suchfunktion (Produkte, die dem Schlagwort des Kunden entsprechen, werden gefunden und angezeigt), *Existence Search* hingegen verlangt eher nach einer übersichtlichen Produktkategorisierung sowie Filter- und Sortierfunktionen. Kategorienübergreifende Vergleichsmöglichkeiten können an dieser Stelle hilfreich sein. Die Suchfunktion sollte möglichst prominent und in angemessener Größe platziert werden [vgl. BVDW 2008(a):14, Sauer, Schmeißer 2005:3], zu vermeiden sind vor allem mögliche Filtereinstellungen, die zu keinem Suchergebnis führen (Preiskategorien, Produkteigenschaften), sowie unverständliche oder falsche Suchergebnisse [vgl. Puscher 2009:123ff.]. Eine dynamisch generierte Autovervollständigung während der Eingabe des Suchbegriffes sowie die Anzeige auch ähnlicher Treffer bei der Suche wirkt sich positiv auf das Nutzungsempfinden aus. Eine Facettenklassifikation der Suchergebnisse sollte ermöglicht werden (bspw. nach Preis, Hersteller, Bewertungen etc.), Bilder sollten angemessen groß und/oder auf Wunsch vergrößerbar sein. Die Produktdarstellungen und -informationen sind optimalerweise einheitlich, so dass eine Orientierung des Kunden gewährleistet ist, und beinhalten Informationen zu Verfügbarkeit und Versand. Lange Response-Zeiten von Servern (bspw. bei Flash-Applikationen) sollten vermieden oder dem Nutzer durch einen Fortschrittsbalken deutlich visualisiert werden. Merkmaltzettel und Wunschlisten wirken sich positiv auf die Intuitivität eines Online-Shops aus [vgl. BVDW 2008(b):7ff.], auch das sogenannte Tagging (Verschlagwortung von Objekten durch Nutzer) kann hier einen Mehrwert schaffen [s. BVDW 2008(b):13].

Beim Kauf des gewünschten Objekts sind vor allem Möglichkeiten von Hinzufügen/Entfernen von Produkten zum Warenkorb sowie die Länge und Abwicklung des Bestellvorgangs von Relevanz

[vgl. Bargas-Avila 2001:24]. Der Warenkorb sollte von jeder Seite aus erreichbar sein [s. BVDW 2008(a):7ff.], dem Nutzer sollte stets angezeigt werden, welche Produkte sich im Warenkorb befinden, Hinweise zu Produktrückgabe, Lieferzeiten und Versandkosten sollten deutlich sichtbar platziert sein [vgl. BVDW 2008(b):14, Sartzetaki et al. 2003:4, Sauer, Schmeißer 2005:6]. Im konkreten Interaktionsprozess mit der Website ist vor allem die Formulargestaltung wichtig. Der Einsatz von sogenannten Reset-Buttons für Eingabefelder kann sich als problematisch erweisen (irrtümliche Löschung von Eingaben). Auch die Wiederherstellbarkeit von Daten (bspw. bei Betätigung des „Zurück“-Buttons in Formularfeldern oder im Vorfeld getätigte Filter-Einstellungen) ist ein wichtiges Kriterium der Gebrauchstauglichkeit in Online-Shops [vgl. Puscher 2009:124, BVDW 2008(b):18]. Bestellvorgänge sollten möglichst kurz gehalten werden, um die Abbruchquote zu vermindern, eine Fortschrittsanzeige im Bestellvorgang kann auch hier zur Nutzerzufriedenheit beitragen [vgl. Sauer, Schmeißer 2005:6].

Zusätzliche Faktoren für einen erfolgreichen E-Commerce Webauftritt sind vor allem ein interessant gestalteter Ablauf sowie das Schaffen einer Vertrauensbasis für den Nutzer (Bewertungssysteme, Gütesiegel, Testberichte) [vgl. Puscher 2009:128]. Raganathan und Ganapathy weisen darauf hin, dass Aspekte wie (Daten-)Sicherheit und Datenschutz im E-Commerce ebenfalls Einfluss auf das Kaufverhalten und die Zufriedenheit der Kunden haben [vgl. Raganathan, Ganapathy 2002:462].

#### ERMITTELTE USABILITY-FAKTOREN:

Kriterium	Autor
Kaufprozess des Kunden möglichst realitätsgetreu abgebildet	Bargas-Avila 2001
Interessant gestalteter Ablauf	Puscher 2009
Übersichtlichkeit von Startseite und Hauptnavigation	Puscher 2009, BVDW 2008(a), Sauer, Schmeißer 2005
Navigationsleiste auf der linken Seite übersichtlich und klar strukturiert	Puscher 2009
Übersichtliche Produktkategorisierung	Puscher 2009, BVDW 2008(a)
Begrenzte Anzahl von Kategorien	Puscher 2009, BVDW 2008(a)
Kategorienübergreifende Vergleichsmöglichkeiten	Sauer, Schmeißer 2005
Effiziente Suchfunktion	Puscher 2009, BVDW 2008(a) Sauer, Schmeißer 2005
Suchfunktion möglichst prominent und angemessen groß	BVDW 2008(a), Sauer, Schmeißer 2005
Autovervollständigung während der Eingabe des Suchbegriffes	BVDW 2008(a)

Keine unverständlichen oder falschen Suchergebnisse	Puscher 2009
Anzeige auch ähnlicher Treffer bei der Suche	BVDW 2008(a)
Facettenklassifikation der Suchergebnisse	BVDW 2008(a)
Filter- und Sortierfunktionen	Puscher 2009
Keine Filtereinstellungen, die zu keinem Suchergebnis führen	Puscher 2009
Produktdarstellungen und -informationen einheitlich	BVDW 2008(a)
Informationen zu Verfügbarkeit und Versand bei Produktinformationen	BVDW 2008(a)
Bilder angemessen groß und/oder auf Wunsch vergrößerbar	BVDW 2008(a)
Fortschrittsbalken bei Ladezeiten	BVDW 2008(a)
Keine langen Response-Zeiten	BVDW 2008(a)
Tagging von Produkten	BVDW 2008(b)
Merkzettel und Wunschlisten	BVDW 2008(a)
Problemloses Hinzufügen/Entfernen von Produkten im Warenkorb	Bargas-Avila 2001
Der Warenkorb von jeder Seite aus erreichbar	BVDW 2008(a)
Ständige Anzeige der im Warenkorb befindlichen Artikel	BVDW 2008(a), Sartzetaki et al. 2003, Sauer, Schmeißer 2005
Hinweise zu Produktrückgabe, Lieferzeiten und Versandkosten im Warenkorb	BVDW 2008(a), Sartzetaki et al. 2003, Sauer, Schmeißer 2005
Länge und Abwicklung des Bestellvorgangs begrenzen	Sauer, Schmeißer 2005, Bargas-Avila 2001
Fortschrittsanzeige im Bestellvorgang	Sauer, Schmeißer 2005
Wiederherstellbarkeit von Daten bei Vor- und Zurücknavigation	Puscher 2009, BVDW 2008(a)
Reset-Buttons für Eingabefelder vermeiden	Puscher 2009
Vertrauensbasis für den Nutzer schaffen	Puscher 2009
Verkäufer/Nutzer – Bewertungssysteme	Puscher 2009
Bewertung des Shops durch Drittanbieter	Puscher 2009
(Daten-)Sicherheit und Datenschutz	Raganathan, Ganapathy 2002

Tabelle 7: Usability-Kriterien der Kategorie E-Commerce

#### 7.4.3 ENTERTAINMENT

Auch der Bereich des Entertainments verlagert sich zunehmend in den Online-Kontext. Anbieter wie [www.lastfm.de](http://www.lastfm.de) (Internetradio) oder [www.youtube.com](http://www.youtube.com) (Videostreams) verzeichnen steigende Nutzerzahlen, 36% der deutschen Bevölkerung nutzen Online-Radio, 27% verwenden Internet-TV

[s. Schneller 2009:35]. Neben rein medienbasierten Seiten wie Radio oder TV fallen auch sogenannte Edutainment-Seiten, (Education + Entertainment, spielerische Vermittlung von Lerninhalten), interaktive Elemente wie Browser- bzw. Online-Games, sogenannte Virtual Reality Chats oder sonstige Unterhaltungselemente, die meist über Plugins in den Browser eingebunden werden können (bspw. Online-Mischen von eigener Musik) in die genannte Kategorie. Neben rein funktionalen Elementen (Browserkompatibilität, Systemanforderungen) stehen vor allem visuelle Aspekte wie Design (Zielgruppenspezifität, Ästhetik) und die Originalität der vermittelten Inhalte im Vordergrund [vgl. Desurvire et al. 2007:3ff.]. Wichtig ist hierbei die Erfüllung der Nutzererwartung bezüglich der Interaktion mit der Website: Es muss sich ein Mehrwert durch Neuheitswert, Originalität und Unterhaltung ergeben. Die Website sollte den Nutzer zum explorativen Surfen animieren, wobei ausreichend Inhalt geboten sein muss, um den Besucher für einen längeren Zeitraum zu unterhalten. Vermieden werden sollte vor allem der Eindruck, dass die Seite bereits nach kurzem Zeitraum keinen Neuheits- oder Mehrwert mehr zu bieten hat. Die Nutzungsmöglichkeiten sollten klar verständlich sein (Downloads, Videos oder interaktive Elemente sollten klar als solche gekennzeichnet sein) [vgl. Wiberg 2005:6f.]. Pinharez et al. (2001) beurteilen Web-Entertainment ebenfalls nach den Kriterien *entertaining*, *engaging*, *appealing* sowie *interactivity* [vgl. Pinharez et al. 2001:12], auch Mandryk et al. (2006) führen die Faktoren *interaction*, *communication*, *engaging experience* und *fun* an [s. Mandryk et al. 2006:142] und evaluieren Computerspiele unter den Gesichtspunkten *boring*, *challenging easy*, *engaging*, *exciting*, *frustration* und *fun* [a.a.O. :2006:152]. Desurvire und Wiberg (2009) haben eine Heuristik zur Evaluation von Spielen entwickelt, von denen einige Kriterien auch auf Browsergames oder Unterhaltung im allgemeinen angewendet werden können. Auch Kirste et al. (2002) beschäftigen sich mit der Übertragbarkeit von Usability-Aspekten in Computerspielen auf andere Benutzeroberflächen. Die drei Autoren weisen darauf hin, dass unterschiedliche Schwierigkeitslevel angeboten werden sollten, so dass der Nutzer längerfristig Freude an der Benutzung empfindet. Sich wiederholender Inhalt sollte möglichst vermieden werden, eine Vielzahl unterschiedlicher Aktivitäten trägt zur längerfristigen Bindung des Nutzers bei. Der Nutzer sollte jederzeit Kontrolle über die angebotenen Funktionen haben, übertragen auf Entertainment-Websites fallen unter diesen Gesichtspunkt vor allem das Abspielen und Anhalten von Multimedia-Elementen. Ausführliche Tutorials (die prominent platziert sind), oder sog. „Guided Tours“ durch das System erleichtern den Einstieg in die Software [vgl. Desurvire, Wiberg 2009:5ff., Kirste et al. 2002:44].

## ERMITTELTE USABILITY-KRITERIEN:

Kriterium	Autor
Nutzungsmöglichkeiten klar verständlich	Wiberg 2005
Zielgruppenspezifisches Design	Desurvire et al. 2007, Pinharez et al. 2001
Mehrwert durch Neuheitswert, Originalität und Unterhaltung	Wiberg 2005, Pinharez et al. 2001
Zum explorativen Surfen animierend	Wiberg C. 2005
Ausreichend Inhalt um den Besucher für einen längeren Zeitraum zu unterhalten	Wiberg 2005
Originalität	Desurvire et al. 2007
Interaktivität	Pinharez et al. 2001, Mandryk et al. 2006
Kommunikationsmöglichkeiten	Mandryk et al. 2006
Herausforderung	Mandryk et al. 2006
Frustration	Mandryk et al. 2006
Spaß	Mandryk et al. 2006, Pinharez et al. 2001
Vielzahl unterschiedlicher Aktivitäten	Desurvivre, Wiberg 2009, Kirste et al. 2002
Unterschiedliche Schwierigkeitsgrade	Desurvivre, Wiberg 2009, Kirste et al. 2002
Nutzerkontrolle über Funktionen	Kirste et al. 2002
Vermeidung des Eindrucks, dass die Seite bereits nach kurzem Zeitraum keinen Neuheits- oder Mehrwert mehr zu bieten hat	Wiberg 2005
Downloads, Videos oder interaktive Elemente klar als solche gekennzeichnet	Wiberg 2005
Tutorials und Anleitungen / Guided Tour vorhanden und prominent platziert	Desurvivre, Wiberg 2009, Kirste et al. 2002
Browserkompatibilität	Desurvire et al. 2007
Systemanforderungen	Desurvire et al. 2007

Tabelle 8: Usability-Kriterien der Kategorie Entertainment

### 7.4.4 INFORMATION

Das World Wide Web hat sich mittlerweile zu einem der bedeutendsten Informationsmedien für die Nutzer entwickelt. 55% der 14- bis 69-jährigen in Deutschland nutzen es, um sich über (Tages-)Politik zu informieren, 47% nutzen das Internet, um über aktuelle wirtschaftliche Entwicklungen auf dem Laufenden gehalten zu werden [s. Schneller 2009:7f.], 41% der deutschen Bevölkerung beziehen Newsletter. Eine wichtige Rolle spielen dabei Online-Auftritte von Zeitungen (www.sueddeutsche.de, www.spiegel-online.de) oder Anbieter wie www.yahoo.de oder www.web.de, die neben einer Suchmaschinenfunktion ebenfalls aktuellen Informationsgehalt unterschiedlichster Art für den Nutzer bereitstellen. Auch Online-Nachschlagewerke



([www.wikipedia.de](http://www.wikipedia.de), [www.wissen.de](http://www.wissen.de)) oder -kataloge von Bibliotheken (OPAC's) können in diese Kategorie gefasst werden. Poll legt in ihrer Arbeit zur Evaluationen von Online-Bibliothekskatalogen dar, dass ein Nutzer entweder gezielt auf der Suche nach einer bestimmten Information ist, oder eine bestimmte Tätigkeit auf der Seite anstrebt will. Übertragen auf den allgemeinen Kontext von informativen Seiten entspricht das einer gezielten Informationsrecherche (Beispiel: Wetter in Hildesheim am 24.12.2009) oder einem allgemeinen Informationsbedürfnis (Beispiel: Aktuelle Tagespolitik) [s. Poll 2007:2]. Vor allem bei Suchmaschinen ist hierbei die Simplizität der Benutzeroberfläche von besonderer Relevanz, um dem Besucher der Website die Informationsrecherche zu erleichtern [vgl. Collins 2006:3]. Artikel o.ä. sollten über entsprechend über eine Freitextsuche auf der Website abrufbar sein, wenn sie spezieller Natur sind oder vom Nutzer nicht durch die Seitenstruktur gefunden werden können. Die Autovervollständigung von Suchbegriffen kann an dieser Stelle hilfreich sein. Unverständliche Begriffe sollten durch die Existenz eines Glossars geklärt werden. Die allgemeine Struktur von Information sollte sich vom Allgemeinen zum Speziellen bewegen (bspw. durch Abstracts, Zusammenfassungen) und erkennbar gegliedert sein (Zwischenüberschriften, ein Gedanke pro Absatz in textueller Information) [vgl. Ludin 2008:4, Harms et al. 2002:284ff.]. Harms et al. (2002) weisen in ihrer Arbeit zur Methodenforschung in der Web Usability darauf hin, dass bei Seiten, die speziell der Erfüllung eines Informationsbedürfnisses des Nutzers dienen, zum einen die Aktualität, zum anderen die Verständlichkeit der Information (zielgruppenorientierte Formulierung) von Relevanz ist. Gelieferte Information sollte vollständig, korrekt, datiert, zuordenbar (Name und Position des Autors), neutral sowie frei von Rechtschreibfehlern sein, um die Glaubwürdigkeit der jeweiligen Quelle zu stützen [vgl. Harms et al. 2002:284ff., Stvilia et al. 2005:2f., Hammwöhner 2007:13ff.]. Moturu und Liu (2009) stellen die Bedeutung der Vertrauenswürdigkeit von Wikipedia-Artikeln, basierend auf Qualität und Glaubwürdigkeit in ihrer Arbeit in den Vordergrund [vgl. Moturu, Liu 2009:1f.]. Collins (2006) führt in diesem Zusammenhang an, dass Werbebanner oder Pop-ups möglichst vermieden werden beziehungsweise begrenzt eingesetzt werden sollten [vgl. Collins 2006:4]. Mariage and Vanderdonckt (2000) weisen in ihrer Studie zur Usability von E-Papers darauf hin, dass der Textumfang einer Webseite nicht mehr als eineinhalb Seiten Text beinhalten sollte. Wichtige Informationen sollten innerhalb des auf Anhieb lesbaren Bereiches des Bildschirms sein [vgl. Mariage, Vanderdonckt 2000:326ff.]. Kurze Sätze und eine einfache Wortwahl unterstützen Nutzer, Informationen schnell und ohne größeren Aufwand zu verarbeiten. Ein Link zur Startseite sollte auf jeder Seite vorhanden und prominent platziert sein [s. Ludin 2008:4]. Kriterien wie Aktualität, Verständlichkeit oder Textumfang sind Usability-Anforderungen, die im allgemeinen zu den generischen Usability-Kriterien gerechnet werden sollen, sind aber im Kontext von Websites, die speziell der Bereitstellung von Information für den Nutzer dienen, von besonderer Relevanz, und werdend daher in der vorliegenden Arbeit mit berücksichtigt, soweit notwendig.

## ERMITTELTE USABILITY-KRITERIEN:

Kriterium	Autor
Simplizität der Benutzeroberfläche	Collins 2006
Freitextsuche	Ludin 2008, Harms et al. 2002
Autovervollständigung von Suchbegriffen	Harms et al. 2002
Abstracts, Zusammenfassungen	Ludin 2008, Harms et al. 2002
Erkennbare Gliederung von Information (Zwischenüberschriften, ein Gedanke pro Absatz in textueller Information)	Harms et al. 2002
Struktur von Informationen vom Allgemeinen zum Speziellen	Harms et al. 2002
Information vollständig, korrekt, datiert, zuordenbar (Name und Position des Autors), neutral sowie frei von Rechtschreibfehlern	Harms et al. 2002, Stvilia et al. 2005f., Hammwöhner R. 2007
Aktualität und Verständlichkeit	Harms et al. 2002
Textumfang einer Webseite nicht mehr als eineinhalb Seiten Text	Mariage, Vanderdonckt 2000.
Wichtige Informationen innerhalb des auf Anhieb lesbaren Bereiches	Mariage, Vanderdonckt 2000
Kurze Sätze und einfache Wortwahl	Ludin 2008, Harms et al. 2002
Zielgruppenorientierte Formulierung	Harms et al. 2002
Qualität und Glaubwürdigkeit	Moturu, Liu 2009
Existenz eines Glossars	Harms et al. 2002
Vermeidung von Werbebannern oder Pop-ups	Collins 2006

Tabelle 9: Usability-Kriterien der Kategorie Information

### 7.4.5 PORTALE, COMMUNITIES, SOCIAL SOFTWARE

Mit Trends aus den USA wie Facebook oder Twitter, die in Deutschland ebenfalls weitgehend verbreitet sind wurden, hat sich der Bereich der Social Software und Communities fest etabliert und stellt mittlerweile im Social Media Marketing oder beim sogenannten Business Networking einen wichtigen Bereich im Web dar. Portale werden in der vorliegenden Arbeit als Internetauftritte mit Kommunikationsfunktion, auf der sich Nutzer unter Erstellung eines Profils anmelden können aufgefasst. Kennzeichnend für solche Webauftritte mit Portalfunktion ist – neben einer einmaligen Anmeldung unter Angabe von persönlichen Daten – vor allem die Möglichkeit zur Erstellung eines in unterschiedlichen Grade detaillierten Nutzerprofiles sowie die Möglichkeit, über eine Vielzahl an Funktionen mit anderen Nutzern zu kommunizieren und zu interagieren. Portale können eine rein kommunikative Funktion (Blogs, Foren) spezifischer (bspw.

Ernährung, Lifestyle) oder allgemeiner Natur haben. Online Communities mit einem klaren Fokus auf Selbstrepräsentation werden aktuell immer populärer ([www.facebook.com](http://www.facebook.com), [www.studivz.net](http://www.studivz.net)) und liefern neben kommunikativen Funktionen auch präsentative (Selbstbeschreibung, Aussagen über Vorlieben und Hobbies) und identifikative (Zugehörigkeit zu Gruppen) Faktoren. Solche sogenannte Social Software kann wiederum dem reinen Selbstzweck (Vernetzung mit Freunden und Bekannten) dienen oder eine spezielle Ausrichtung wie haben (E-Recruiting: [www.xing.de](http://www.xing.de), [www.monster.com](http://www.monster.com)).

Ein wichtiger Faktor, der bei der Konzeption und dem Design von Online-Portalen vor allem beachtet werden sollte, ist die Kommunikation zwischen den Nutzern. Beiträge sollten möglichst leicht und unkompliziert zu veröffentlichen sowie durch Tags annotierbar oder durch andere Personen zu editieren sein [vgl. Koch 2008:2]. Die Funktionen sollten deutlich sichtbar und leicht zu bedienen sein, bestehende Formatierungsvorlagen (Schriftschnitt, Einbindung von Grafiken) können den Umgang für unerfahrene Nutzer erleichtern. Analogien zu bekannten Programmen (bspw. die Einbindung einer an Microsoft Office Word ausgerichteten Menüleiste) können die Erwartungskonformität der Anwendung an dieser Stelle fördern [vgl. Müller-Prothmann, Siedentopf 2003:4, Preece 2004:4f.]. Helic et al. (2004) verweisen im Zusammenhang mit Online-Foren vor allem auf die Relevanz einer vereinfachten Benutzeroberfläche [vgl. Helic et al. 2004:12]. Die Ausrichtung der Community sollte klar und deutlich kommuniziert werden (Sinn, Funktionen), die Plattform sollte möglichst rund um die Uhr erreichbar sein, da Nutzungsprobleme aufgrund von Wartungsarbeiten oder Serverausfällen zur in Abschnitt 2.1 beschriebenen Abwanderung von Nutzern führen können. Funktionen zur Profilerstellung sollten vorhanden (Schaffung von Identität und Vertrauensbildung), deutlich sichtbar und einfach zu bedienen sein (bspw. sollte der Ausschluss bestimmter Bildformate beim Hochladen eines Profilbildes oder eine Größenbegrenzung sichtbar angezeigt werden). Die Einrichtung eines Belohnungsmodus (Mitglieder-Ranking, Bewertungsmöglichkeiten) für sinnvolle oder nützliche Beiträge kann vor allem in sogenannten Support-Foren (bspw. für Soft- und Hardware) Anreize und positives Erleben für die Nutzer schaffen [vgl. Preece 2004:4f., Helic et al. 2004:13]. Für neue Mitglieder sollte – vor allem bei Portalen mit komplexen Strukturen – eine Einführung oder eine Tour durch das System angeboten werden [vgl. Preece 2004:4f.]. Das Design des Layouts und der angebotenen Informationen (die nicht durch Nutzer generiert und veröffentlicht wurden) sollten möglichst zielgruppenspezifisch umgesetzt werden. Datenschutzrechtliche Aspekte sind vor allem in diesem Kontext wichtig, da die Nutzer zum Teil höchst private Daten in der Community veröffentlichen. Hinweise auf rechtliche Auswirkungen und Umgang mit privaten Informationen sollten entsprechend prominent gekennzeichnet werden [s. Müller-Prothmann, Siedentopf 2003:4].

## ERMITTELTE USABILITY-KRITERIEN:

Kriterium	Autor
Ausrichtung der Community klar und deutlich kommuniziert	Preece 2004
Design des Layouts und der angebotenen Informationen möglichst zielgruppenspezifisch	Müller-Prothmann, Siedentopf 2003
Vereinfachte Benutzeroberfläche	Helic et al. 2004
Funktionen deutlich sichtbar und leicht zu bedienen	Müller-Prothmann, Siedentopf 2003, Preece 2004
Funktionen zur Profilerstellung vorhanden (Schaffung von Identität und Vertrauensbildung), deutlich sichtbar und einfach zu bedienen	Preece 2004
Ausschluss bestimmter Bildformate (beim Hochladen eines Profilbildes) oder Größenbegrenzung von Dateien sichtbar angezeigt	Preece 2004
Beiträge sollten möglichst leicht und unkompliziert zu veröffentlichen	Koch 2008.
Formatierungsvorlagen (Kursivschnitt, Einbindung von sog. Emoticons)	Müller-Prothmann, Siedentopf 2003, Preece 2004
Analogien zu bekannten Programmen (Einbindung einer an Microsoft Office Word ausgerichteten Menüleiste)	Müller-Prothmann, Siedentopf 2003
Editierbarkeit von Beiträgen	Koch 2008.
Tagging	Koch 2008.
Belohnungsmodus	Preece 2004, Helic et al. 2004
Tour durch das System	Preece 2004
Hinweise auf rechtliche Auswirkungen und Umgang mit privaten Informationen	Müller-Prothmann, Siedentopf 2003
Plattform möglichst rund um die Uhr erreichbar	Preece 2004

Tabelle 10: Usability-Kriterien der Kategorie Social Software

#### 7.4.6 PRÄSENTATION

Die Repräsentation von (Privat-)Personen, Organisationen, Firmen und Einrichtungen im Internet ist mittlerweile zu einem gängigen Standard geworden. Der Nutzer soll auf privaten oder kommerziellen Homepages über den Anbieter informiert werden. Werbung in eigener Sache und eine ansprechende Darstellung ist so für Firmen mit begrenztem Aufwand möglich. Rein repräsentative Seiten (bspw. von Banken, Organisationen, Privatpersonen), wie sie im vorherigen Abschnitt von informativen Seiten abgegrenzt wurden, werden mittlerweile immer mehr von Websites, die dem Nutzer einen interaktiven Mehrwert (Online-Banking, Download von Informationsmaterialien oder Tools) bieten, abgelöst. Dennoch stellen sie einen wichtigen Teilbereich des World Wide Web dar. Zu beachten bei der Erstellung von repräsentativen Homepages sind ähnliche Kriterien wie bei der Erstellung informativer Seiten. Lange Ladezeiten von Bildern sollten möglichst vermieden werden, bei der Einbindung von Plugins oder Applets sollte auf möglichst geringe Systemanforderungen geachtet werden, um den größtmöglichen Nutzerkreis ansprechen zu können. Bilder und Plugins sollten Alternativtexte besitzen, um sogenannte „Plain-Text-Browser“ im Zuge der Barrierefreiheit zu unterstützen. Das Profil des Anbieters sollte umfassend und in ausreichender Tiefe dargestellt werden, wobei auf die bereits im vorangegangenen Abschnitt geschilderten Richtlinien, was Darstellung, Farbwahl und Schrift betrifft, geachtet werden sollte [Bevan 1998:2ff.]. Redundante Informationen sollten vermieden werden [vgl. Nielsen 2001<sup>19</sup>]. Erkenntlich sollte vor allem werden, was den Anbieter der Homepage auszeichnet, beziehungsweise welchen Nutzen der Besucher in Abgrenzung zu anderen Anbietern oder Websites hat. Kontaktdaten sollten deutlich sichtbar und leicht über die Navigation zu finden sein [vgl. Nielsen 2003<sup>20</sup>]. Newsletter oder RSS-Feeds können eine längerfristige Bindung des Nutzers bewirken. Auch hier sind zielgruppenspezifische (Terminologie, Mehrsprachigkeit der Website) Aspekte zu beachten. Sitemaps können helfen, dem Nutzer schnell und effektiv einen Überblick über Struktur und Inhalte der Seite zu verschaffen [vgl. Bevan 1998:2ff.].

#### ERMITTELTE USABILITY-KRITERIEN:

Kriterium	Autor
Profil des Anbieters umfassend, aktuell und in ausreichender Tiefe dargestellt	Bevan 1998
Erkennbarkeit, was den Anbieter der Homepage auszeichnet / welchen Nutzen der Besucher in Abgrenzung zu anderen Anbietern oder Websites hat	Nielsen 2003
Kontaktdaten deutlich sichtbar und leicht über die Navigation zu finden	Nielsen 2003

<sup>19</sup> <http://www.useit.com/homepageusability/guidelines.html>, Zugriff 27.03.2010

<sup>20</sup> <http://www.useit.com/alertbox/20031110.html>, Zugriff 27.03.2010

Keine redundanten Informationen	Nielsen 2001
Textumfang einer Webseite nicht mehr als eineinhalb Seiten	Mariage, Vanderdonckt 2000
Wichtige Informationen innerhalb des auf Antrieb lesbaren Bereiches	Mariage, Vanderdonckt 2000
Struktur von Informationen vom Allgemeinen zum Speziellen	Harms et al. 2002
Erkennbare Gliederung (Zwischenüberschriften, ein Gedanke pro Absatz) des Textes	Harms et al. 2002
Zielgruppenspezifische Terminologie	Bevan 1998
Keine langen Ladezeiten von Bildern	Bevan 1998
Plugins oder Applets stellen geringe Systemanforderungen	Bevan 1998
Alternativtexte für Bilder und Plugins	Bevan 1998
Sitemaps	Bevan 1998
Existenz eines Glossars	Harms et al. 2002
Newsletter oder RSS-Feeds	Bevan 1998
Keine Werbebanner oder Pop-ups	Collins 2006

Tabelle 11: Usability-Kriterien der Kategorie Repräsentation

#### 7.4.7 E-LEARNING

Auch der Bereich der Vermittlung von Lerninhalten wird von Unternehmen und Einrichtungen in zunehmendem Maße in den Online-Kontext verlagert. Nutzer von E-Learning-Software, die meist zum Download bereitgestellt wird oder direkt auf der Website des Anbieters verfügbar ist, besitzen oft über wenig oder kein Vorwissen im Umgang mit Software. Usability-Aspekte sind hier entsprechend von besonderer Relevanz. Usability im E-Learning-Kontext beinhaltet zwei Faktoren: pädagogische Gebrauchstauglichkeit (angemessene Vermittlung der dargebotenen Lerninhalte) sowie technische Gebrauchstauglichkeit (einfacher, problemfreier Umgang mit dem System), mit dem Ziel, den Aufwand an kognitiven Ressourcen im Umgang mit dem System zu minimieren und so einen maximalen Lerneffekt mit Freude des Nutzes am Gebrauch (*joy of use*) zu erreichen [vgl. Kukulska-Hulme 2004:1, Melis, Weber 2003:2, Zaharias 2007:1]. Personalisierung spielt im E-Learning eine wichtige Rolle. Die angebotenen Inhalte sollten positiv mit dem Lernprozess des Nutzers korrelieren (je länger der Lernprozess des Nutzers, desto höher das Niveau der Inhalte), wobei die grafische Benutzeroberfläche jedoch konstant bleiben sollte, um Desorientierung zu vermeiden [vgl. Melis, Weber 2003:2f.]. Eine großzügige Anordnung der einzelnen Elemente, die ausreichend Außenabstand lässt, kann sich hierbei positiv auf das Nutzungsempfinden auswirken. Auch motorische Fehler können so vermieden werden (Reduzierung der Anzahl „falscher“ Klicks durch zu kleine oder zu nah platzierte Schaltflächen). Die Differenzierung des

Schwierigkeitsgrades sollte eher in größeren Clustern gehalten sein (einfach – mittel – schwer), eine entsprechende farbliche Gestaltung der grafischen Benutzeroberfläche schafft Orientierung des Benutzers. Icons sollten einfach gehalten sein, und keine komplexen Funktionen beinhalten [vgl. Heidenreich, Raff 2006:2ff.]. Die verwendete Terminologie sollte zielgruppenspezifisch angepasst sein, abstrakte oder komplizierte Konzepte mithilfe grafischer Abbildungen veranschaulicht werden [s. Zaharias 2007:7]. Hilfe- und Supportfunktionen sollten differenziert (unterschiedliche Typen, beispielsweise eine Tour durch das System zusammen mit ausführlicher Dokumentation und FAQ) und deutlich sichtbar sowie jederzeit verfügbar sein. Online-Foren und Wikis können den gegenseitigen Austausch der Lernenden und eigene Problemlösungen fördern [vgl. Rentría-Bonito et al. 2006:187], der ergänzende Einsatz von Multimedia-Elementen (Videos, Audiodateien) kann zur Veranschaulichung der Lehrinhalte nützlich sein [vgl. Latten 2001:5]. Der Nutzer sollte nach Möglichkeit in der Lage sein, frei zu navigieren, also die Option haben, innerhalb unterschiedlicher Lernmodule zu wechseln [s. Zaharias 2007:7]. Zur Hilfefunktion ist im Allgemeinen zu unterscheiden zwischen pädagogischer (Hilfestellung zu Übungsaufgaben, Hinweise zu Inhalten, Lösungsmöglichkeiten) und technischer (Installation des Systems, Inhaltliche Struktur der Inhalte) Hilfefunktion bei der Nutzung. Aus pädagogischer Sicht sollten vor allem bei Hilfe zu Lösungsmöglichkeiten nur Hinweise, nicht jedoch sofort die komplette Lösung angeboten werden. Erklärungen bei Falschantworten sollten geliefert werden [vgl. Melis, Weber 2003:3, Katre 2007:63]. E-Learning-Nutzer im Beginner-Stadium benötigen eher aktive Hilfestellung (automatische Hinweise des Systems), bereits erfahrenen Lernern hingegen sollte es ermöglicht werden, Hilfefunktionen nur auf Anfrage präsentiert zu bekommen. Richtlinien zur Textgestaltung auf dem Bildschirm wie in Abschnitt 7.4.6 und 7.4.4 geschildert, sollten auch hier beachtet werden. Druckfunktionen für lange Texte sollten vorhanden sein, um dem Lernenden zu ermöglichen, Inhalte via Papier aufzuarbeiten [vgl. Melis, Weber 2003:3ff.].

#### ERMITTELTE USABILITY – KRITERIEN:

Kriterium	Autor
Konstante grafische Benutzeroberfläche	Melis, Weber 2003
Großzügige Anordnung der einzelnen Elemente	Heidenreich, Raff 2006
Farbliche Gestaltung der grafischen Benutzeroberfläche	Heidenreich, Raff 2006
Einfache Icons ohne komplexe Funktionen	Heidenreich, Raff 2006
Freie Navigierbarkeit (Option, innerhalb unterschiedlicher Lernmodule zu wechseln)	Zaharias 2007
Terminologie zielgruppenspezifisch angepasst	Zaharias 2007
Angebotene Inhalte entsprechen dem Lernfortschritt des Nutzers	Melis, Weber 2003
Differenzierung des Schwierigkeitsgrades in größeren Stufen	Heidenreich, Raff 2006
Abstrakte oder komplizierte Konzepte mithilfe	Zaharias 2007

grafischer Abbildungen veranschaulicht	
Ergänzender Einsatz von Multimedia-Elementen (Videos, Audiodateien) zur Veranschaulichung der Lehrinhalte	Latten 2001
Online-Foren und Wikis	Rentróia-Bonito et al. 2006
Hilfe zu Lösungsmöglichkeiten: Hinweise statt kompletter Lösung	Melis, Weber 2003, Katre 2007
Erklärungen bei Falschantworten	Melis, Weber 2003, Katre 2007
Hilfe- und Supportfunktionen differenziert (unterschiedliche Typen, Systemtour, Dokumentation, FAQ) deutlich sichtbar sowie jederzeit verfügbar	Rentróia-Bonito et al. 2006
Sowohl aktive als auch passive Hilfestellung	Melis, Weber 2003
Druckfunktionen für lange Texte	Melis, Weber 2003

**Tabelle 12: Usability-Kriterien der Kategorie E-Learning**

## 7.5 ZUSAMMENFASSUNG

Es lässt sich festhalten, dass für alle der in Abschnitt 7.2 ermittelten Website-Kategorien einschlägige Usability-Kriterien ermittelt werden konnten, die zur späteren Generierung von kategoriespezifischen Fragebogenitems dienen sollen. Hierbei lag die Kategorie E-Commerce mit den meisten Publikationen zum Thema Usability und bezüglich übereinstimmender Aussagen zu den Usability-Kriterien deutlich vorn. Wünschenswert wären mehr Publikationen in den Bereichen Social Software und Entertainment gewesen, da in diesen Kategorien die Usability-Forschung zum Teil bisher noch in der Entwicklung begriffen ist. Vor allem im Bereich Entertainment sind Publikationen meist auf ein bestimmtes Thema spezialisiert, die Ableitung allgemeiner Kriterien für die Gebrauchstauglichkeit wird eher selten vorgenommen. Auch für den Bereich Information sind allgemeine Richtlinien und Hinweise in der Literatur eher begrenzt vorhanden. An diesen Stellen wäre weitere Forschung auf jeden Fall sinnvoll und wünschenswert. Die kategoriespezifischen Usability-Kriterien wurden dennoch soweit wie möglich ermittelt, und sollen im folgenden Abschnitt geordnet und strukturiert werden.

## 7.6 EINORDNUNG DER ERMITTELTEN KRITERIEN

Zur Strukturierung der Kriterien waren mehrere Optionen möglich. Beispielsweise wäre eine Unterteilung in die in Abschnitt 2.3 beschriebenen Dimensionen der Web-Usability (Design, Content, Struktur) möglich gewesen, oder eine dichotome Unterteilung in Content und Struktur. Die Entscheidung, die ermittelten Kriterien nach den sieben Grundsätzen der Dialoggestaltung nach ISO 9241 zu gliedern ergab sich aus der Fragestellung der Arbeit (Beurteilung der



Gebrauchstauglichkeit) und aus der für die Fragestellung passenden, wissenschaftlich anerkannten Struktur, die durch die Grundsätze der ISO-Norm bereitgestellt wird.

Nachdem die Entscheidung für eine Strukturierungsweise der ermittelten Usability-Kriterien gefallen war, wurden im folgenden Schritt die Kriterien den Grundsätzen der Dialoggestaltung zugeordnet. Hierbei wurde für jedes Kriterium, das in den vorangegangenen, kategoriespezifischen Abschnitten festgehalten wurde, der passendste Gestaltungsgrundsatz ausgewählt. Die Zuordnung wurde durch den Autor vorgenommen, und ist somit das Ergebnis einer subjektiven Einschätzung. Teilweise wäre die Zuordnung zu unterschiedlichen Kategorien möglich gewesen, in diesen Fällen wurde die Kategorie beruhend auf der für den Autor vorherrschenden Schwerpunktsetzung innerhalb eines Kriteriums ausgewählt. Für den zu erstellenden Fragebogengenerator ist in erster Linie die pragmatische Strukturierung der Kriterien an sich von Bedeutung. Da die Fragestellung, welchem von zwei oder mehr passenden Gestaltungsgrundsätzen ein Kriterium grundsätzlich zuzuordnen ist, eine eher sekundäre Rolle spielt, ist dieses Vorgehen in der vorliegenden Arbeit zu rechtfertigen.

Kriterien, die augenscheinlich keinem der sieben Grundsätze der Dialoggestaltung zuzuordnen waren, wurden unter dem Stichpunkt „ergänzender Kriterienkatalog“ zusammengefasst.

Anmerkung zur Kategorie Entertainment:

In der Gestaltung von Websites mit dem Schwerpunkt Unterhaltung kann neben den Grundsätzen der Dialoggestaltung ebenfalls auf die ISO EN 14915 (Software-Ergonomie für Multimedia-Benutzungsschnittstellen) zurückgegriffen werden. Die Iso-Norm nennt folgende Grundsätze zur Gestaltung von Multimedia-Anwendungen:

- Eignung für das Kommunikationsziel
- Eignung für Wahrnehmung und Verständnis
- Eignung für Exploration
- Eignung für Benutzungsmotivation

[<http://www.handbuch-usability.de/iso-14915.html>]

Entsprechend wurden Items der Kategorie Multimedia zusätzlich zur Einordnung in die allgemeinen Grundsätzen der Dialoggestaltung den Kriterien der ISO 14915 zugeordnet, soweit nötig.

Bei der Strukturierung der Kriterien wurden diese zum Teil wörtlich aus den Tabellen 6-11, zum Teil sinngemäß aus dem Fließtext übernommen. Zum Teil wurden Kriterien aus der Literatur

weiter unterteilt<sup>21</sup>. Eine wortgenaue Entsprechung ist also nicht in allen der zugeordneten Kriterien gegeben.

Nachfolgend findet sich eine beispielhafte Zuordnung der ermittelten Usability-Kriterien des Bereichs Entertainment zu den einzelnen Gestaltungsgrundsätzen. Die restlichen Tabellen finden sich entsprechend in Appendix B.

Grundsätze der Dialoggestaltung	Zugeordnete Kriterien
Aufgabenangemessenheit	<ul style="list-style-type: none"> <li>▪ Ausreichend Inhalt zur Unterhaltung des Besuchers für einen längeren Zeitraum vorhanden</li> <li>▪ Systemanforderungen</li> <li>▪ Browserkompatibilität</li> <li>▪ Originalität der Inhalte</li> </ul>
Selbstbeschreibungsfähigkeit	<ul style="list-style-type: none"> <li>▪ Nutzungsmöglichkeiten klar verständlich</li> <li>▪ Funktionen prominent platziert</li> <li>▪ Icons simpel und eindeutig gehalten</li> </ul>
Steuerbarkeit	<ul style="list-style-type: none"> <li>▪ Entfernung von Plugins möglich</li> <li>▪ Getätigte Einstellungen (Filter etc.) können rückgängig gemacht werden</li> <li>▪ Manueller Start von Multimedia-Daten, Unterbrechung möglich</li> <li>▪ Nutzerkontrolle über die Funktionen</li> </ul>
Erwartungskonformität	<ul style="list-style-type: none"> <li>▪ Design von Tools in Anlehnung an mentale Modelle</li> <li>▪ Zielgruppenspezifisches Design (Alter, Geschlecht)</li> </ul>
Fehlertoleranz	<ul style="list-style-type: none"> <li>▪ Tippfehlertoleranz von Suchfeldern</li> </ul>
Individualisierbarkeit	<ul style="list-style-type: none"> <li>▪ Anpassung des Layouts</li> <li>▪ Empfehlungen von Musikstücken, Videos etc.</li> <li>▪ Größe von Elementen (Videobildschirm, Radioregler, Bildschirmgröße) anpassbar</li> </ul>
Lernförderlichkeit	<ul style="list-style-type: none"> <li>▪ Guided Tour oder interaktive Systemhilfe vorhanden</li> <li>▪ Tutorials oder Anleitungen vorhanden und sichtbar platziert</li> </ul>

**Tabelle 13: Zuordnung beispielhafter Usability-Kriterien zu den Grundsätzen der Dialoggestaltung**

Grundsätze der Gestaltung von Multimedia-Anwendungen	
Eignung für das Kommunikationsziel	<ul style="list-style-type: none"> <li>▪ Deutliche Kennzeichnung von interaktiven Elementen (Downloads, Funktionsbuttons)</li> <li>▪ Ausreichende Kommunikationsmöglichkeiten zwischen Benutzern</li> </ul>
Eignung für Wahrnehmung und Verständnis	<ul style="list-style-type: none"> <li>▪ Layout dem Entertainment-Typus angepasst</li> <li>▪ Größe des Videobildschirms ausreichend</li> <li>▪ Platzierung von Elementen auf der Seite</li> <li>▪ Interaktivitätsgrad der Seite</li> </ul>

<sup>21</sup> Beispielsweise enthält der Punkt „Nutzungsmöglichkeiten klar verständlich“ aus nachvollziehbaren Gründen die Kriterien „Funktionen als solche gekennzeichnet“ und „Funktionen auffindbar“.

Eignung für Exploration	<ul style="list-style-type: none"> <li>▪ Zum explorativen Surfen animierend</li> <li>▪ Vielzahl unterschiedlicher Aktivitäten</li> </ul>
Eignung für Benutzungsmotivation	<ul style="list-style-type: none"> <li>▪ Neuheits-, Originalitäts- und Unterhaltungswert der Seite</li> <li>▪ Herausforderung</li> <li>▪ Vorhandensein unterschiedlicher Schwierigkeitsgrade</li> <li>▪ „Spaß“ - wert der Seite</li> </ul>

**Tabelle 14: Zuordnung beispielhafter Usability-Kriterien zu den Grundsätzen der Gestaltung von Multimedia-Anwendungen**

## 7.7 FAZIT

Im vorangegangenen Abschnitt wurde zunächst ein eigenes Klassifikationsschema zur Kategorisierung unterschiedlicher Website-Typen vorgenommen. Im Anschluss konnten spezifische Usability-Kriterien für die einzelnen Kategorien aus der Literatur recherchiert und strukturiert werden. Für die einzelnen Kategorien wurde im großen und ganzen eine zufriedenstellende Anzahl an Usability-Kriterien gefunden, obwohl anzumerken ist, dass vor allem in den Bereichen Social Software und Entertainment weitere Arbeit in der Usability-Forschung wünschenswert wäre. Die Kriterien wurden im Anschluss mit dem Ziel der Übersichtlichkeit und Strukturierung entsprechend den Grundsätzen der Dialoggestaltung zugeordnet. Im folgenden Abschnitt werden aus den ermittelten Kriterien Items für das zu erstellende Fragebogentool generiert werden, die im Anschluss zur qualitativen Absicherung durch Experten und Studierende der Informationswissenschaft getestet werden sollen.

## 8 ERSTELLUNG DER FRAGEBOGENITEMS

Bei der Erstellung der Fragebogenitems lassen sich unter methodischem Gesichtspunkt folgende zwei Probleme identifizieren: Wahl des Itemtyps und Wahl der Skala. Im Abschnitt 8.1 sollen die Entscheidungen, die zur jeweiligen Wahl geführt haben, erklärt und begründet sowie die Erstellung der Items beschrieben werden. Test und Überarbeitung der Items werden entsprechend in Abschnitt 8.2 bis Abschnitt 8.7 erläutert.

### 8.1 METHODISCHES VORGEHEN ZU SKALENWAHL UND ITEM TYP

Ziel des Fragebogens, der mit dem späteren Fragebogentool generiert werden soll, ist es, Einstellungen und Ansichten von Benutzern zur Usability einer Website zu erheben. Zum Itemtyp fiel die Wahl daher auf eine fünfstufige Likert-Skala, die durch Zustimmung oder Ablehnung zu einer Aussage die Einstellung eines Benutzers zu einem bestimmten Sachverhalt, einem Produkt oder Objekt erheben kann [vgl. Bortz, Döring 2006:252]. In allen in Abschnitt 6 vorgestellten Fragebögen zur Erhebung von Web- oder Software-Usability werden Aussage-Items oder semantische Differentiale, welche Aussagen entsprechen, verwendet. Dies trug, zusammen mit den Ergebnissen der sozialwissenschaftlichen Literaturrecherche im Bereich

Fragebogenkonstruktion (s. Abschnitt 5.1) dazu bei, bei dem zu wählenden Item auf den Typ „Aussage-Item“ zurückzugreifen.

Die Anzahl der Stufen der Likert-Skala wurde anhand von eigenen Überlegungen zur Übersichtlichkeit auf dem Bildschirm bei verschiedenen Auflösungen und Monitorgrößen getroffen. Die Wahrscheinlichkeit für den Benutzer, bei sieben oder mehr Stufen horizontal scrollen zu müssen, ist zum einen höher als bei fünf Stufen. Der Fragebogen soll zum anderen für den Benutzer auf den ersten Blick einen übersichtlichen und ansprechenden Eindruck machen, um die Non-Response-Quote bereits im Vorfeld zu senken [s. Sarodnick, Brau 2006:172]. Unter pragmatischen Gesichtspunkten lässt sich an dieser Stelle ebenfalls die Frage stellen, ob der Spezialisierungsgrad einer Skala mit sieben oder mehr Stufen dem zu erhebenden Sachverhalt überhaupt angemessen ist. Ob ein Nutzer, der (ohne weitergehende Vorbereitung) am Ende eines Besuchs auf einer Website mit einem Kurzfragebogen zur Usability konfrontiert wird, überhaupt in der Lage ist, seine Einstellung zur Gebrauchstauglichkeit einer Website in einem derart hochgradig untergliederten Skalenverhältnis kundzutun, ist fraglich. Entsprechend fiel die Wahl auf eine fünfstufige Skala.

Bei der Erstellung der Items wurde jeweils nach der Reihenfolge der Grundsätze der Dialoggestaltung vorgegangen. Die einzelnen Kriterien aus Abschnitt 7.6 wurden zu Aussagen, denen der Nutzer mit Zustimmung oder Ablehnung begegnen kann, formuliert. Kriterien, die durch Experten im Rahmen einer heuristischen Evaluation abgefragt werden könnten, wurden entweder ausgelassen, oder im Rahmen desselben Sachverhalts umformuliert<sup>22</sup>. Aussagen, die dichotom (ja/nein) zu beantworten wären, wurden nach Möglichkeit vermieden. Unter Berücksichtigung der Formulierungsregeln der empirischen Sozialforschung bei der Fragebogenkonstruktion, wie in Abschnitt 5.1 beschrieben, wurden negative Formulierungen vermieden, um den kognitiven Aufwand für den Nutzer so gering wie möglich zu halten. Die Formulierungen beschränken sich nach Möglichkeit auf einfache Sätze, die jeweils ein Merkmal oder einen Sachverhalt abfragen.

Zur Wertigkeit der Aussagen lässt sich festhalten, dass darauf verzichtet wurde, dasselbe Merkmal sowohl mit negativer als auch mit positiver Wertigkeit abzufragen obwohl dies, wie in Abschnitt 5.1 beschrieben, zu empfehlen wäre. Entsprechend wurden die meisten der Items positiv formuliert. Da sich das Medium World Wide Web jedoch, wie beschrieben, durch eine hohe Nutzerfluktuation und deren erhöhte Abbruchbereitschaft bei Umfragen auszeichnet, muss, um eine möglichst hohe Teilnehmerzahl erreichen zu können, die Priorität auf eine begrenzte Länge des Fragebogens gelegt werden. Vor dem Hintergrund der pragmatischen Ausrichtung des Fragebogens lässt sich vertreten, dass unter sozialwissenschaftlichen Gesichtspunkten möglicherweise eine geringfügige Beeinflussung der Antwortergebnisse durch die Itemformulierung vorliegt. Da die mit dem Fragebogen arbeitenden Nutzer keine repräsentative Stichprobe für eine Grundgesamtheit darstellen, sondern konkrete Hinweise auf

---

<sup>22</sup> Beispiel: Kriterium „Sichtbarkeit des Warenkorbs auf jeder Seite“ wurde umformuliert zu Item „Der Warenkorb ist leicht zu finden“.

Nutzungsprobleme und Usability-Probleme liefern sollen, ist in dieses Vorgehen in der vorliegenden Arbeit nach Ansicht des Autors zu rechtfertigen. Insgesamt wurden 179 Items generiert. Die Aufteilung auf die unterschiedlichen Kategorien waren hierbei wie folgt:

Kategorie	Anzahl Items
E-Commerce	49
Entertainment	31
E-Learning	31
Social Software	25
Information	26
Repräsentation	17

Tabelle 15: Anzahl an Items in den jeweiligen Kategorien

Die nachfolgende Tabelle stellt einige Formulierungen von Beispielitems aus dem Bereich E-Commerce dar. Die kompletten Itemsätze finden sich in Appendix C.

Item	Itemformulierung
<b>Beispielitems:</b>	
	„Der Warenkorb ist leicht auf der Seite zu finden.“
	„Die Suchfunktion lieferte mir nützliche Ergebnisse.“
	„Die Anordnung der Produkte auf der Seite verwirrt mich.“
<b>Vermiedene Formulierung von Items:</b>	
Aussage dichotom	„Ein Warenkorb ist immer sichtbar.“
Aussage von Experten zu beurteilen	„Ein Warenkorb existiert.“
Aussage zu komplex (mehrere Merkmale)	„Der Warenkorb ist leicht zu finden und immer sichtbar.“
Aussage zu lang	„Wenn ich die Seite besuche oder auf der Seite einkaufe gehe, kann ich den Warenkorb immer leicht finden.“
Aussage negativ formuliert	„Der Warenkorb ist nicht leicht für mich zu finden.“
Aussage doppelt wertig abgefragt	„Der Warenkorb ist leicht zu finden.“
	„Der Warenkorb ist schwer für mich zu finden.“

Tabelle 16: Beispielhafte Itemformulierungen

Nachdem die Fragebogenitems generiert worden waren, mussten diese im nächsten Schritt auf Validität (tatsächliche Messung des zu erfassenden Sachverhalts) überprüft werden. Die Items wurden dabei auf die Kriterien Relevanz, Vollständigkeit sowie Verständlichkeit geprüft. Dies geschah zum einen durch eine Experten-Revision, zum anderen durch eine Revision der Items durch Studierende der Informationswissenschaft. Anschließend wurden die Itempools für die jeweiligen Kategorien gemäß den Hinweisen beider Gruppen optimiert.

## 8.2 REVISION DES ITEMPOOLS DURCH EXPERTEN

Insgesamt wurden zwölf Experten gebeten, die Items gemäß ihrer Expertise und Erfahrung im Bereich Usability zu beurteilen. Experten stellten in diesem Fall Universitätsmitarbeiter und berufstätige Personen im Bereich Usability dar. Ebenfalls wurden Absolventen der Informationswissenschaft mit Abschlussarbeiten im entsprechenden Themenbereich befragt. Das Erkenntnisinteresse bei dieser ersten Art der Revision lag vor allem auf Relevanz der Items einer Kategorie für den jeweiligen Usability-Bereich. Hierzu wurde eine dichotome Bewertungsmöglichkeit vorgegeben (relevant/nicht relevant, s. Abbildung 1). Der zweite Schwerpunkt der Experten-einschätzung lag auf der Beurteilung der Vollständigkeit des Itempools. Mit diesem Kriterium sollte vor allem sichergestellt werden, dass die endgültigen Fragebogenitems den zu evaluierenden Sachverhalt (Usability der Website) vollständig erfassen (Inhaltsvalidität, vgl. Abschnitt 5.2.1.3).

Zu diesem Zweck wurden die jeweiligen generierten Itempools in tabellarische Form gebracht, sowie mit einem Anschreiben und einer Erklärung zur Beurteilung der Items versehen (s. Appendix D). Jedes Item beinhaltete die Möglichkeit zur Abgabe eines Freitextkommentars.

Aufgabenangemessenheit	besonders relevant	nicht relevant
Der Warenkorb ist leicht auf der Seite zu finden	<input type="checkbox"/>	<input type="checkbox"/>
Das Vorhandensein eines Merkzettels erleichtert mir meinen Einkauf	<input type="checkbox"/>	<input type="checkbox"/>
Ich begrüße die Möglichkeit, Waren und Produkte nach ihren Eigenschaften (Preis, Relevanz) sortieren zu können	<input type="checkbox"/>	<input type="checkbox"/>
Die Anordnung der Produkte auf der Seite verwirrt mich	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 11: Ausschnitt aus einem Experten-Testbogen

Weiterhin wurden die Experten gebeten, Angaben zu Alter, Geschlecht und beruflichen Erfahrungen und/oder Kenntnissen im Bereich Usability zu machen. Dies sollte dazu beitragen, die Ergebnisse der Beurteilung sowie Änderungen im Itempool qualitativ abzusichern und intersubjektiv überprüfbar zu machen. Die Experten beurteilten – je nach Expertise – zwischen

einem und sechs Bögen<sup>23</sup>. Insgesamt konnte so eine Anzahl von 29 Bewertungen für Kategorien erreicht werden, die sich folgendermaßen aufteilte:

Kategorie	Expertenurteile
<b>E-Commerce</b>	10
<b>Entertainment</b>	3
<b>E-Learning</b>	5
<b>Social Software</b>	8
<b>Information</b>	4
<b>Repräsentation</b>	7

**Tabelle 17: Anzahl Expertenurteile für die jeweiligen Kategorien**

Die klare Mehrheit der Beurteilungen im Bereich E-Commerce lässt sich an dieser Stelle durch eine entsprechende berufliche Tätigkeit der Experten erklären, und stellt nicht unbedingt einen methodischen Nachteil, sondern eher einen qualitativen Vorteil bezüglich der Güte der Bewertung für diese Itemkategorie dar. Weitere Bewertungen im Bereich Entertainment sowie Information wären jedoch entsprechend wünschenswert. Aus Gründen des begrenzten zeitlichen Umfangs der Arbeit musste an dieser Stelle allerdings auf ein weiteres Anschreiben von Experten verzichtet werden, da die Items nach Abschluss der Expertenbeurteilungen zusätzlich einem Test mit einer größeren Stichprobe unterzogen werden sollten.

### 8.3 REVISION DES ITEMPOOLS DURCH STUDIERENDE DER INFORMATIONSWISSENSCHAFT

Zusätzlich zur Beurteilung der Fragebogenitems durch Experten wurde parallel eine Revision der Items durch Studierende der Informationswissenschaft vorgenommen. Diese wurde mit Seminarteilnehmern durchgeführt. Schwerpunkt des Erkenntnisinteresses war an dieser Stelle vor allem die Verständlichkeit der Formulierungen. Das Design der Items war analog zu demjenigen der Experten (tabellarische Form) mit Anpassungen im Anschreiben (siehe Appendix D). Zusätzlich wurden in die zu bearbeitenden Bögen „Fehler“ (Items anderer Websitekategorien) eingebaut, um zu überprüfen, ob die Aussagen als kategoriefremd eingestuft würden. Die Studierenden wurden mit der Bitte angeschrieben, sich in die Rolle eines Benutzers ohne Usability-Vorwissen zu versetzen und Unklarheiten oder Unverständlichkeiten zu kommentieren. Weiterhin wurden die

<sup>23</sup> Die Bögen mit den jeweiligen Itempools aus den Kategorien wurden als .doc – Format verschickt, auf Wunsch wäre auch eine Beurteilung der Bögen als Papierversion möglich gewesen. Diese Möglichkeit wurde nicht in Anspruch genommen.

Studierenden ebenfalls um eine Einschätzung der Relevanz der einzelnen Items gebeten<sup>24</sup>. Diese sollte vor allem zum Abgleich mit der Einschätzung der Experten dienen.

Insgesamt konnte so eine Anzahl von 84 Bewertungen erreicht werden<sup>25</sup>, die sich folgendermaßen aufteilte:

Kategorie	Bewertungen
E-Commerce	13
Entertainment	15
E-Learning	14
Social Software	14
Information	13
Repräsentation	15

Tabelle 18: Anzahl Studierendenurteile für die jeweiligen Kategorien

#### 8.4 RELEVANZERRECHNUNG

Nachdem alle Experten- und Nutzerurteile eingegangen waren, wurde im Anschluss ein Relevanzwert für jedes Item jeder Kategorie berechnet. Dies sollte zum einen dazu dienen, Items, die sowohl von Experten als auch von den Studenten als irrelevant beurteilt wurden, aus dem Itempool zu entfernen. Da inhaltlich ähnliche und doppelte Items von der Mehrheit der Personen als „nicht relevant“ markiert worden waren, sollte dieses Verfahren zum anderen dazu dienen, inhaltsgleiche Items auszusortieren.

Die Items sollen im später zu erstellenden Tool als Orientierungshilfe für den Nutzer eine Kennzeichnung aufweisen, die es ermöglichen soll, besonders wichtige Items herauszustellen. Da der Nutzer die Items für seinen Fragebogen selbst zusammenstellen kann, soll an dieser Stelle eine intuitive Hilfestellung zur Itemwahl geliefert werden. An dieser Stelle soll auf den hier errechneten Relevanzwert zurückgegriffen werden. Items, die von Experten und Studenten als hoch relevant bewertet worden waren, können entsprechend visuell hervorgehoben werden.

Hierzu wurde im ersten Schritt getrennt nach Experten und Studenten für jedes Item der Relevanzwert  $R_{\text{Item}}$ , basierend auf dem Verhältnis von „relevant“-Bewertungen zur Gesamtzahl der Bewertungen errechnet. Entsprechend war

$$R_{\text{Item}} = \frac{\text{Anzahl relevant-Bewertungen}}{\text{Gesamtzahl der Bewertungen}}$$

mit  $0 < R_{\text{Item}} < 1$ .

<sup>24</sup> Die Bearbeitung der Bögen fand zum einen während der Seminarzeit auf Papierbögen und zum anderen als Hausaufgabe elektronisch im .doc-Format statt.

<sup>25</sup> Insgesamt nahmen 40 Studierende an dieser Aufgabe teil. 22 Teilnehmer füllten jeweils drei Bogen zuhause als Papierversion aus, 18 Bögen wurden direkt im Seminar ausgefüllt.



Hierbei ergaben sich folgende gerundete Mittelwerte<sup>26</sup> für die Relevanzscores in den jeweiligen Kategorien:

Kategorie	Mittelwert $R_{\text{ItemE}}$ <sup>27</sup> Experten	Mittelwert $R_{\text{ItemS}}$ <sup>28</sup> Studierende	Abweichung $\Delta R_{\text{ItemE}} - R_{\text{ItemS}}$
E-Commerce	0,7854	0,7676	0,0178
Entertainment	0,7419	0,757	0,0151
E-Learning	0,7677	0,8594	0,0917
Social Software	0,695	0,6943	0,0007
Information	0,7212	0,7604	0,0392
Repräsentation	0,7983	0,7843	0,014
Mittelwert $\sum R_{\text{Item}}$	0,7515	0,7705	0,0297

Tabelle 19: Mittelwerte für die Relevanzscores

Die einzelnen Werte für jedes Item können in Appendix E zusammen mit den Freitextkommentaren gefunden werden.

Stärkere Abweichungen zwischen den Relevanzscores der Experten und der Studierenden ergaben sich vor allem im Bereich E-Learning ( $\sim 0,091$ ) und Information ( $\sim 0,039$ ), auffällig ist die starke Übereinstimmung für einen vergleichsweise geringeren Relevanzwert im Bereich Social Software, was auf einen stärkeren Überarbeitungsbedarf hinwies. Im Durchschnitt ergab sich ein Mittelwert der Abweichung von  $\sim 0,0297$ . Zu interpretieren ist dies wie folgt: Experten und Benutzer stimmten vor allem im Bereich Social Software darin überein, dass durchschnittlich mehr Items nicht oder weniger relevant waren als in den anderen Kategorien. Im Bereich E-Learning wichen die Ansichten von Experten und Studierenden stärker als in anderen Kategorien voneinander ab, da die Studierenden deutlich mehr Items als relevant einstufen als die Experten. Insgesamt stimmten die Experten und die Studierenden in Urteilen jedoch in großen Teilen überein.

Um zu ermitteln, wie relevant jedes Item letztendlich bewertet werden sollte, war es nötig, eine Kombination aus Experten- und Studierendenurteilen zu finden. Daher wurde im darauffolgenden Schritt der Durchschnitt aus den errechneten Bewertungen von Experten und Studierenden für jedes Item ermittelt. Bei den studentischen Befragten handelt es sich um Teilnehmer unterschiedlichen Kenntnistanandes und Semester. Es kann entsprechend davon ausgegangen werden, dass die Relevanzbewertungen der Experten aufgrund ihrer Expertise, ihres Fachwissens und zum Teil langjähriger beruflicher Erfahrung als stärker ausschlaggebend angesehen werden können. Dementsprechend wurde der ermittelte Relevanzscore der Experten (in der Formel abgekürzt als  $R(E)$ ) bei der Berechnung des Relevanzwertes doppelt so stark gewichtet, wie derjenige der Studierenden (abgekürzt  $R(S)$ ). Der Relevanzwert für die jeweiligen Items ergab sich damit als

<sup>26</sup> Mittelwert über die Relevanzscores aller Items einer Kategorie

<sup>27</sup>  $R_{\text{ItemE}}$ : Experten-Relevanzscore für ein Item

<sup>28</sup>  $R_{\text{ItemS}}$ : Studierenden-Relevanzscore für ein Item

$$R_{\text{Item}} = \frac{2 \cdot R(E) + R(S)}{3}.$$

Nachfolgend dargestellt die Mittelwerte für die berechneten Relevanzbewertungen für die Items der jeweiligen Kategorien:

Kategorie	Mittelwert $R_{\text{Item}}$
E-Commerce	0,7795
Entertainment	0,747
E-Learning	0,7983
Social Software	0,6948
Information	0,7342
Repräsentation	0,7937
Mittelwert $\sum R_{\text{Item}}$	0,7579

**Tabelle 20: Mittelwerte der endgültigen Relevanzscores der einzelnen Kategorien**

Die durchschnittliche Relevanzbewertung der Studierenden und Experten von ~0,76 war als ein zunächst zufriedenstellendes (wurden doch die erstellten Items von der Mehrzahl der Teilnehmer als relevant für die jeweilige Kategorie angesehen), jedoch verbesserungsbedürftiges Ergebnis zu bewerten.

## 8.5 FREITEXTKOMMENTARE DER EXPERTEN

Die Experten wurden mit der Bitte angeschrieben, zusätzlich zu ihrer Einschätzung bezüglich der Relevanz Kommentare zu den einzelnen Items abzugeben, falls nötig. Da die Bögen von den Experten und Studierenden zum Teil auf Papier kommentiert worden waren, wurden die Anmerkungen zur besseren Operationalisierbarkeit zunächst in eine Excel-Tabelle übertragen. Die kompletten Ergebnisse finden sich in Appendix E. Angemerkt wurden vor allem Redundanzen, die Spezifiziertheit einzelner Items (zu spezifisch/zu ungenau) sowie inhaltliche Aspekte. Vermehrt kamen Hinweise auf die globale (kategorienübergreifende) Gültigkeit von Items (z.B. Hilfe- und Suchfunktion).

Nachfolgend ein beispielhafter Auszug aus den Kommentaren der Experten für den Bereich E-Commerce und E-Learning<sup>29</sup>:

<sup>29</sup> aus Visualisierungsgründen wurde die Struktur der ursprünglichen Tabelle unter Beibehaltung des Inhalts geändert.

Item	Experte I <sup>30</sup>	Experte II	Experte III
1) „Die Hilfefunktion konnte meine Fragen klären.“	„Hier könnte man differenzieren in FAQ, Index, Forum... Relevant.“	„Finde ich nicht so gut formuliert. Jmd., der keine Fragen hatte, fühlt sich nicht angesprochen. Relevant.“	„Relevant.“
2) „Wenn ich ein bestimmtes Produkt suche, kann ich es leicht über Suche und Navigation finden.“	„Besser nur nach Suchfunktion fragen: „Die Suche brachte mir nützliche Ergebnisse.“ Nicht Relevant.“	„Relevant.“	„Suche und Navigation getrennt abfragen[...] Auf Navigation geht aber der darüber liegende Punkt m.E. auch schon ausreichend ein. Relevant.“
3) „Es fällt mir leicht, mich auf den Lernstoff zu konzentrieren.“	„Bezug zur Website fehlt, besser: „Die Aufmachung der Seite lenkt nicht vom Lernstoff ab.“ Relevant.“	„Der Grund für die Ausprägung der Antwort auf diese Frage kann zusammenhangslos zu den relevanten Fragen sein. Irrelevant.“	„Relevant.“

**Tabelle 21: Auszug der Expertenkommentare für den Bereich E-Commerce und E-Learning**

Zu interpretieren sind diese Angaben am Beispiel der hier aufgeführten Kommentare wie folgt: Die Expertenmeinungen für Aussage 1 unterscheiden sich hinsichtlich des Iteminhalts, bewerten es aber alle als relevant. Die Konsequenz für das methodische Vorgehen impliziert die Überlegung, die Hilfefunktion weiter zu differenzieren und das Anbieten der Option „keine Angabe/nicht zutreffend“ im späteren Bogen für Umfrageteilnehmer, die die Hilfefunktion nicht genutzt haben. Ähnlich verhält es sich mit den zwei anderen Items. Die zweite Aussage sollte inhaltlich zweigeteilt werden, um jeweils nur eine der genannten Funktionen (Suche und Navigation) abzufragen. Die dritte Aussage sollte gemäß der Anregung von Experte I umformuliert werden, da Experte II das berechtigte Argument vorbrachte, dass unter Verwendung der ursprünglichen Wortwahl der abgefragte Sachverhalt unter Umständen keinen Zusammenhang mit der zu evaluierende Website aufweisen würde.

<sup>30</sup> Experte I, II und III seien beispielhafte Personen. Es handelt sich nicht bei allen vier Items um jeweils dieselben Experten, vielmehr wurden für jedes Item Experten mit beispielhaften Anmerkungen ausgewählt.

## 8.6 FREITEXTKOMMENTARE DER STUDENTEN

Die Studenten waren im Anschreiben gebeten worden, sich in die Rolle eines Benutzers ohne Vorwissen im Bereich Usability hineinzusetzen. Ziel der Befragung war vor allem die Prüfung auf Verständlichkeit der Aussagenformulierungen. Neben vermehrten Hinweisen auf inhaltlich redundante Items wurden vor allem Anglizismen sowie Fachausdrücke angemerkt. Auch das Problem zu komplexer Itemformulierungen wurde des Öfteren genannt. Die kategoriefremden Items, die in die Bögen eingebaut worden waren, wurden von den Studierenden allesamt gefunden. Dies kann als Gütekriterium für die Inhaltsvalidität der Items angesehen werden, da die Items den jeweiligen Bereichen offensichtlich klar und unstrittig zuzuordnen waren.

Beispielhaft sei an dieser Stelle der nachfolgende Auszug aus den Studierendenkommentaren für den Bereich E-Commerce und Entertainment<sup>31</sup>:

Item	Studierender I <sup>32</sup>	Studierender II	Studierender III
1) „Wenn ich ein bestimmtes Produkt suche kann ich es leicht über Suche und Navigation finden.“	„doppelt. Relevant“	„Suchfunktion statt Suche. Relevant.“	„Relevant.“
2) „Die Anordnung der Produkte nach Tägung einer Filtereinstellung (z.B. Sortierung nach Preis – aufsteigend) ist mir unerklärlich.“	„Formulierung! Besser: Anordnung der Produkte nach vorheriger Filterung. Relevant.“	„Nicht relevant.“	„Besser: unverständlich. Relevant“
3) „Tutorials und Anleitungen sind so platziert, dass ich sie leicht finde.“	„Begriffsklärung Tutorial.“	„Begriff Tutorial nicht klar.“	„Relevant.“

**Tabelle 22: Auszug der Studierendenkommentare für den Bereich E-Commerce und Entertainment**

Für das erste Item lässt sich der Schluss ziehen, dass für den Studierenden I die Aussage inhaltlich nicht klar genug abgegrenzt von anderen Fragen zu Navigation war (da diese bereits in vorhergehenden Items abgefragt wurde), und so inhaltlich doppelt erscheint. Alle drei Studierenden bewerteten das Item jedoch als relevant. Das zweite Item sollte hinsichtlich der Formulierung überdacht werden, da zwei Hinweise diesbezüglich kamen. Analog verhält es sich mit der dritten Aussage, wo der Begriff Tutorial offensichtlich durch einen anderen Begriff ersetzt oder mit einer Erklärung versehen werden sollte.

<sup>31</sup> aus Visualisierungsgründen wurde die Struktur der ursprünglichen Tabelle unter Beibehaltung des Inhalts geändert

<sup>32</sup> Studierender I, II und III seien beispielhafte Personen. Es handelt sich nicht bei allen Items um jeweils dieselben Studierenden, vielmehr wurden für jedes Item Studierende mit beispielhaften Anmerkungen ausgewählt.

## 8.7 ÜBERARBEITUNG DES ITEMPOOLS

Nachdem die komplette Auswertung der Experten- und Studierendenkommentare erfolgt war, wurde der gesamte Itempool überarbeitet und Aussagen mit geringer Relevanzbewertung entfernt, sowie Anregungen gemäß der Freitextkommentare eingearbeitet. Die Kommentare beider Gruppen wurden hier gleichermaßen berücksichtigt, da für die spätere Anwendbarkeit des Bogens sowohl die inhaltliche Genauigkeit als auch die angemessene Formulierung der Items eine wichtige Rolle spielt. Vorgenommen wurde vor allem die Umformulierung von Anglizismen und von als unverständlich gekennzeichneten Ausdrücken. Auch die Vereinfachung des Satzbaus war ein wichtiger Teil der Überarbeitung der Aussagen. Nachfolgend ein Beispiel für das Vorgehen bei der Überarbeitung des Items „*Wenn ich ein bestimmtes Produkt suche kann ich es leicht über Suche und Navigation finden.*“:

- 1) Vergleich der Relevanzeinschätzung von Experten und Studierenden (vier von sechs Aussagen kennzeichnen das Item als relevant, s. Abschnitt 8.5 und 8.6).
- 2) Einarbeitung der Formulierungsvorschläge: Umbenennung von Suche zu Suchfunktion und Aufteilung des Items in zwei Aussagen.
- 3) Überprüfung, ob andere, inhaltsgleiche Items für Suchfunktion oder Navigation vorliegen (nicht der Fall).
- 4) Item I: „Die Hauptnavigation stellt eine gute Orientierungshilfe dar.“
- 5) Weitere Aufteilung der Aussage zur Suchfunktion.
- 6) Item II: „Die Suchfunktion ist leicht zu finden.“
- 7) Item III: „Die Suchfunktion liefert mir nützliche Ergebnisse.“

Neben der Überarbeitung von Items wurde ebenso eine Reduzierung des Itempools vorgenommen. Redundante Aussagen wurden zusammengefasst und Items mit niedrigem Relevanzscore zum größten Teil entfernt. Unter Eliminierung von Items, die vermehrt als redundant gekennzeichnet worden waren, erfolgte die Reduzierung der endgültigen Items auf die folgenden Anzahlen:

Kategorie	Itemzahl vor Bewertung	Itemzahl nach Bewertung
E-Commerce	49	33
Entertainment	31	24
E-Learning	31	28
Social Software	25	23
Information	26	25
Repräsentation	17	16
Σ Items	179	149

Tabelle 23: Anzahl der Items vor und nach der Bewertung durch Experten und Studierende

Die stärksten Kürzungen wurden hierbei im Bereich E-Commerce und Entertainment vorgenommen<sup>33</sup>. Nachfolgend findet sich ein beispielhafter Auszug für die Überarbeitung und/oder Eliminierung von Items aus den Bereichen E-Commerce. Für eine vollständige Darstellung der überarbeiteten Items siehe Appendix F.

Ursprüngliches Item	Überarbeitete Itemversion
„Ich begrüße die Möglichkeit, Waren und Produkte nach ihren Eigenschaften (Preis, Relevanz) sortieren zu können.“	<i>„Ich kann die Produkte meinen Wünschen entsprechend sortieren.“</i>
„Die Hilfefunktion konnte meine Fragen klären	<i>„Die Hilfefunktion ist ausreichend für mich.“</i>
„Manchmal lösche ich Eingaben, weil ich versehentlich auf den Zurücksetzen-Button gekommen bin.“ / „Den „Eingaben zurücksetzen“-Button finde ich überflüssig.“	Eliminiert, da Relevanzscore für beide Items $\leq 0,5$ und Kommentare der Experten entsprechend
„Ich habe mich durch die Bewertung einer unabhängigen Institution oder Einrichtung (z.B. Stiftung Warentest) für den Shop entschieden.“	Eliminiert, da entsprechende Kommentare der Experten

Tabelle 24: Beispielhafter Vergleich von ursprünglichen und überarbeiteten Items

Nach der Überarbeitung der Items konnten – unter Berücksichtigung der zuvor abgegebenen Relevanzwerte – folgende Relevanzwerte erreicht werden:

Kategorien	Relevanzwert nach Überarbeitung
E-Commerce	0,857
Entertainment	0,783
<b>E-Learning</b>	0,813
Social Software	0,778
<b>Information</b>	0,792
Repräsentation	0,805

Die Werte sind in allen Kategorien höher, als die zuvorigen Relevanzwerte. Die Überarbeitung der Items ist also unter dem Gesichtspunkt der Relevanz der Items als sinnvoll zu betrachten.

<sup>33</sup> Ausnahmen wurden bei in der Literatur als besonders relevant beschriebenen Kriterien, die von den Experten als weniger relevant eingestuft wurden, gemacht.

## 8.8 PRETEST DER ITEMS MIT ENDNUTZERN

Vor den endgültigen Tests wurden die überarbeiteten Items der jeweiligen Kategorien weiteren Nutzern zu einem letzten Pretest vorgelegt. Dies sollte dazu dienen, eventuelle Schwierigkeiten in Nutzung und Umgang im Vorfeld aufzudecken, und gegebenenfalls zu beheben. Hierzu wurden die Items zusammen mit der Skala in tabellarische Form gebracht und das Layout demjenigen der späteren Online-Umfrage angepasst<sup>34</sup>:

The image shows a screenshot of a pretest questionnaire. It contains three items, each followed by a five-point Likert scale. The items are:

- Über den Anbieter der Seite werden ausreichend Informationen geliefert
- Die Informationen über den Anbieter der Seite sind auf dem neuesten Stand
- Die Informationen über den Anbieter der Seite sind leicht verständlich

Each item has a corresponding scale with five options: ☐ stimme gar nicht zu, ☐ stimme weniger zu, ☐ teils teils, ☐ stimme eher zu, ☐ stimme voll zu.

Abbildung 12: Ausschnitt aus einem Pretest-Benutzerbogen

Insgesamt wurde jeder der sechs Bogen von zwei Benutzern getestet. Die Arbeitsanweisung lautete, den Bogen anhand von beliebigen Websites der jeweiligen Kategorien auszufüllen<sup>35</sup>. Zwei der Bögen wurden unter gleichzeitiger Anwesenheit von Testleiter und Tester getestet (E-Commerce, E-Learning). Die restlichen zehn der Bögen wurden räumlich entfernt getestet. Von diesen fanden vier Tests synchron mit Livechat-Präsenz<sup>36</sup> statt. Die übrigen Bögen wurden per E-Mail verschickt. Diese Mischung aus zeitgleichen (synchronen) und zeitversetzten (asynchronen) Pretests sollte sicherstellen, dass eventuell aufkommenden Fragen zum einen sofort dokumentiert werden konnten, zum anderen die reale Umgangssituation mit dem Fragebogen optimal simuliert werden konnte.

Sowohl die synchronen als auch die asynchronen Pretests mit endgültigen Nutzern bereiteten keine Probleme. Nachfragen kamen lediglich bezüglich der Nichtexistenz abgefragter Funktionen und Elemente auf. Da den Testern alle Items der jeweiligen Kategorien vorlagen, ließ sich nicht vermeiden, dass zum Teil Funktionalitäten abgefragt wurden, die die evaluierte Seite nicht anbot<sup>37</sup>. Dies verdeutlicht die Notwendigkeit einer „keine Antwort“/„nicht zutreffend“ –

<sup>34</sup> Die Fragebogenitems wurden den Nutzern als Word-Dokument vorgelegt.

<sup>35</sup> Pro Kategorie wurden zwei Beispielwebsites genannt, die jedoch nicht obligatorisch waren. Da es in dieser Phase der Itementwicklung noch nicht auf die Ergebnisse der Beurteilung der Website sondern nur auf die Anwendbarkeit der Items an sich ankam, wurde die Wahl der zu evaluierenden Website weitgehend freigestellt.

<sup>36</sup> Testleiter und Tester im peer-to-peer-Netzwerk, Kategorien: Information, Homepages, Social Software, Entertainment.

<sup>37</sup> z.B. Item 11 in der Kategorie Social Software: „Die Kommentierung oder Verschlagwortung von Beiträgen ist einfach“ - Kommentierung von Nutzerbeiträgen in Foren, wird von [www.facebook.com](http://www.facebook.com) angeboten, von [www.studivz.net](http://www.studivz.net) jedoch nicht

Antwortoption in den endgültigen Tests der Items. Weitere Rückmeldungen bezüglich der Verständlichkeit der Items oder des Bogens kamen nicht. Auf Nachfrage gaben alle sechs Beteiligten an, keine weiteren Probleme mit der Bearbeitung der Bögen gehabt zu haben.

## 8.9 FAZIT

Im vorangegangenen Kapitel konnte unter Berücksichtigung der aus der Literatur ermittelten, kategoriespezifischen Usability-Kriterien Fragebogenitems entwickelt werden. Diese sollen die Grundlage für das zu erstellende Fragbogentool bilden, aus dem der Nutzer sich später einen auf seine Website angepassten Fragebogen zusammenstellen kann. Die erstellten Items wurden im zweiten Schritt von Experten und Studierenden getestet, ein Vorgehen, das sowohl die Qualität und die Inhaltsvalidität der Items sicherstellen sowie deren Verständlichkeit und Akzeptanz bei späteren Nutzern beurteilen helfen sollte. Die Items wurden gemäß den Anregungen beider Gruppen überarbeitet und vor der endgültigen Testanwendung mit Benutzern einem letzten Pretest unterzogen. Als Ergebnis des Pretests kristallisierte sich eine klare Notwendigkeit nach einer Antwortkategorie „keine Angabe/nicht zutreffend“ außerhalb der eigentlichen Skala heraus, die im endgültigen Test Einsatz finden sollte. Im folgenden Abschnitt soll nun der Online-Test der Items beschrieben sowie die statistische Analyse der Rücklaufdaten vorgenommen werden.

## 9 BENUTZERTEST DER FRAGEBÖGEN

Nachdem die endgültigen Items als Word-Dokument von Endnutzern getestet und als problemfrei beurteilt worden waren, konnten die Itemsätze mithilfe des Programmes LimeSurvey als Online-Version erstellt werden. Das Vorgehen, die endgültigen Itemsätze mit einer größeren Stichprobe online zu testen, sollte vor allem eine große Nutzerzahl unter geringerem Zeitaufwand ermöglichen. Da ein Präsenztest unter Anwesenheit von Testleiter und Tester zum einen hochgradig zeitaufwendig (bedingt durch die Anzahl an Kategorien) gewesen wäre, zum anderen keinen erkennbaren methodischen Vorteil gebracht hätte, und die Präsenz-Pretests problemlos verlaufen waren, fiel die Entscheidung auf dieses Vorgehen.

### 9.1 ERSTELLUNG DER FRAGEBÖGEN IN LIME SURVEY

LimeSurvey ist eine Software, mit der Online-Umfragen (genannt Surveys) sowohl erstellt als auch durchgeführt werden können, und die gleichzeitig eine Auswertung für die erhobenen Daten bereitstellt. Da an der Universität Hildesheim eine Lizenz für LimeSurvey vorhanden war, fiel die Wahl auf dieses Tool. Zum endgültigen Test der Items wurde für jede Kategorie eine Umfrage erstellt, die alle Items einer Kategorie enthielt.



## 9.2 ONLINE-PRETEST UND DURCHFÜHRUNG DES ENDGÜLTIGEN BENUTZERTESTS

Nachdem die Items für die Kategorien online erstellt und das Layout der Umfrage angepasst worden war, wurde ein letzter Pretest mit zwei Benutzern durchgeführt, der das problemlose Arbeiten mit den Online-Bögen sicherstellen sollte. In Ergebnis dieses Tests wurde ein Item im Bereich Social Software angepasst<sup>38</sup> und kleinere Veränderungen im Layout (Anpassung der Schriftgröße sowie der Farbwahl) vorgenommen. Ziel der Layoutoptimierung war zum einen die verbesserte Lesbarkeit der Texte am Bildschirm, sowie zum anderen die identische Anzeige der Umfrage in den Browsern Mozilla Firefox und Internet Explorer.

Um den abschließenden Test mit Endnutzern durchzuführen, wurde der Link zur Umfrage per E-Mail an potentielle Teilnehmer verschickt. Genutzt wurden hierbei vor allem soziale Netzwerke<sup>39</sup> sowie E-Mail-Verteiler<sup>40</sup> und private Kontakte. Bei den angeschriebenen Personen handelte es sich um Studierende und Berufstätige aller Fachrichtungen. Der Durchführungszeitraum belief sich auf 19 Tage (10. Februar 2010 – 01. März 2010).

## 9.3 SERVER- UND DNS-PROBLEME MIT LIMESURVEY

Beim Pretest mit der endgültigen Online-Version der Fragebögen traten zwei Probleme auf. Zum einen existierte ein Serverproblem welches verhinderte, dass die abschließende Seite der Umfrage angezeigt wurde. Nach Betätigen des „Umfrage-Absenden“-Buttons erschien stattdessen ein weißer Bildschirm. Zum anderen verursachte der DNS-Server des Rechenzentrums der Universität, auf dem die LimeSurvey-Installation gehostet wurde, Probleme, mit dem Ergebnis, dass der Link, unter dem die Umfrage zu erreichen war, bei einem der zwei Nutzer aus dem Pretest nicht funktionierte.

Diese Probleme blieben bis zum Abschluss der Pretest-Phase bestehen, und konnten nicht behoben werden. Da die Probleme netzwerkbedingt bestanden, keine Zugriffsrechte auf die Netzwerkadministration vorhanden waren, und kein zufriedenstellendes alternatives Tool zur Online-Erstellung des Bogens existierte, wurden die abschließenden Tests des Items trotzdem unter LimeSurvey durchgeführt. Unter den angeschriebenen Teilnehmern kam vermehrt die Rückmeldung, dass der per E-Mail versendete Link nicht funktionierte. Als Konsequenz daraus wurde in der zweiten Hälfte des Befragungszeitraumes eine Weiterleitungsfunktion für den Link von der Plattform [www.usability-toolkit.de](http://www.usability-toolkit.de) gelegt. Dies behob das Zugangsproblem zur Umfrage. Wie sich die beschriebenen Probleme auf die Teilnehmerquote ausgewirkt haben, ist nicht zu ermitteln, dennoch ist anzunehmen, dass ohne die Zugangsschwierigkeiten eine höhere Anzahl an Datensätzen hätte erreicht werden können.

<sup>38</sup> „Die Hilfefunktion ist gut sichtbar platziert.“ wurde zu „Die Hilfefunktion ist leicht zu finden.“ umformuliert.

<sup>39</sup> [www.studivz.net](http://www.studivz.net), [www.facebook.de](http://www.facebook.de), [www.lokalisten.de](http://www.lokalisten.de)

<sup>40</sup> studiengangsinterner Verteiler [iplusalle], Verteiler der Stipendiaten der Friedrich-Ebert-Stiftung

## 9.4 DURCHFÜHRUNG DER UMFRAGE

Die Items wurde in den Online-Fragebögen um ein Anschreiben ergänzt, in dem die Nutzer gebeten wurden, sich auf eine bestimmte, der Kategorie zuzuordnenden Website zu begeben, und sich einen kurzen Überblick über die jeweilige Seite zu verschaffen und den Fragebogen auszufüllen. Diese Website wurde im Anschreiben vorgegeben.

Der Begrüßungstext wurde je nach Websitekategorie unterschiedlich verfasst. Neben einem Hinweis auf die zeitliche Dauer der Umfrage wurde auch darauf hingewiesen, dass unter Umständen Funktionen oder Elemente der Seite zur Beurteilung abgefragt werden könnten, die die Seite nicht anbietet. Die Nutzer wurden gebeten, diese Fragen mit der Option „keine Aussage/nicht zutreffend“ zu bewerten. Weiterhin wurde ein Vermerk zum Datenschutz am Anfang der Umfrage platziert.

Anzumerken ist an dieser Stelle, dass das Nichtvorkommen abgefragter Merkmale auf den Seiten einen methodischen Nachteil mit sich bringt, da die Erhebung statistischer Kennzahlen im Nachhinein Einschränkungen unterliegt (siehe Abschnitt 10). Jedoch konnten aus Gründen des zeitlich begrenzten Umfangs der Arbeit nicht zu jeder Kategorie mehrere Websites online evaluiert werden. Somit stellt diese Art der Herangehensweise eine grundlegende statistische Beurteilung der Items dar und lässt die Möglichkeiten weiterer Validierung offen.

**Evaluation: Entertainment-Website**

Hallo!  
Vielen Dank, dass Sie sich Zeit nehmen, den **Webauftritt von youtube.com** zu beurteilen!

Bitte begeben Sie sich auf die Seite [www.youtube.com](http://www.youtube.com). Versuchen Sie, sich einen kurzen Überblick über die Möglichkeiten und Funktionen der Seite zu verschaffen, und beantworten Sie dann den Fragebogen, indem Sie unten auf "Weiter" klicken. Die Dauer sollte in etwa **10 Minuten** betragen.

Für Funktionen, die Sie nicht genutzt haben, oder die die Seite nicht anbietet, klicken Sie bitte "keine Antwort" an.

Unter allen Teilnehmern wird am Schluss ein Gutschein von amazon.de in Höhe von 15,- verlost. Wenn sie an der Verlosung teilnehmen möchten, hinterlassen sie bitte Ihre E-Mailadresse am Ende des Fragebogens.

**Ihre Daten werden selbstverständlich vertraulich behandelt und nicht an Dritte weitergegeben!**

Diese Umfrage enthält 26 Fragen.

**Eine Bemerkung zum Datenschutz**  
Dies ist eine anonyme Umfrage.

Die Daten mit Ihren Antworten enthalten keinerlei auf Sie zurückzuführende/identifizierende Informationen, es sei denn bestimmte Fragen haben Sie explizit danach gefragt. Wenn Sie für diese Umfrage einen Zugangsschlüssel benutzt haben, so können Sie sicher sein, dass der Zugangsschlüssel nicht zusammen mit den Daten abgespeichert wurde. Er wird in einer getrennten Datenbank aufbewahrt und nur aktualisiert, um zu speichern, ob Sie diese Umfrage abgeschlossen haben oder nicht. Es gibt keinen Weg die Zugangsschlüssel mit den Umfrageergebnissen zusammenzuführen.

**Weiter >>**

Abbildung 13: Startseite der Umfrage mit Anschreiben und Datenschutzerklärung

Die Präsentation der Items erfolgte in Gruppen von 5 Fragen pro Bildschirm, was die Lesbarkeit erhöhen und Scrollen auf der Mehrzahl der Monitore vermeiden sollte.

**Evaluation: Entertainment-Website**

0%  100%

**Ich finde mich auf der Seite auf Anhieb zurecht**

Stimme überhaupt nicht zu	Stimme weniger zu	Neutral	Stimme eher zu	Stimme vollkommen zu	nicht zutreffend/keine Antwort
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Die Funktionen der Seite sind auf Anhieb ersichtlich**

Stimme überhaupt nicht zu	Stimme weniger zu	Neutral	Stimme eher zu	Stimme vollkommen zu	nicht zutreffend/keine Antwort
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Die Funktionen der Seite sind einfach zu bedienen**

Stimme überhaupt nicht zu	Stimme weniger zu	Neutral	Stimme eher zu	Stimme vollkommen zu	nicht zutreffend/keine Antwort
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Die Seite bietet auch nach längerer Benutzung noch Unterhaltungswert**

Stimme überhaupt nicht zu	Stimme weniger zu	Neutral	Stimme eher zu	Stimme vollkommen zu	nicht zutreffend/keine Antwort
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Die Seite wirkt originell auf mich**

Stimme überhaupt nicht zu	Stimme weniger zu	Neutral	Stimme eher zu	Stimme vollkommen zu	nicht zutreffend/keine Antwort
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Abbildung 14: Beispielseite der erstellten Umfrage

Die Umfrage unterlag keinen weiteren Beschränkungen, sondern war öffentlich zugänglich. Somit sollte eine größere Anzahl an Teilnehmern erreicht werden. Um sicherzustellen, dass auch alle sechs Fragebögen weitestgehend in gleicher Anzahl beantwortet würden, wurde ein Skript auf den Server gelegt, welches die randomisierte Verteilung der Nutzer auf die einzelnen Fragebögen vornahm. Da sich während der Testphase dennoch eine Ungleichverteilung der Nutzer mit Schwerpunkt auf die drei Kategorien Entertainment, Information und Social Software ergab, erfolgte in der Mitte der Umfragezeit nur noch die Verteilung der Teilnehmer auf die verbleibenden drei Kategorien E-Commerce, Homepages und E-Learning durch Änderung des Skriptes.

Am Ende der Umfrage wurden die Teilnehmer gebeten, bei Interesse an der Verlosung eines 15,- Euro Gutscheines für den Online-Shop [www.amazon.de](http://www.amazon.de) ihre E-Mail-Adresse zu hinterlassen. Ebenfalls wurde ein Freitextfeld für Kommentare oder weitere Anregungen erstellt. Die Hinzufügung weiterer Items fand nicht statt. An der fünfstufigen Likert-Skala fand eine Ergänzung um die Antwortmöglichkeit „nicht zutreffend/keine Angabe“ statt, alle Fragen waren freiwillig auszufüllen. Hiermit sollte sichergestellt werden, dass Features, die auf der Seite nicht existierten oder Items, die der Nutzer nicht beantworten wollte, nicht erzwungenermaßen abgefragt wurden.

**Evaluation: Entertainment-Website**

0%  100%

Bitte geben Sie hier Ihre Emailadresse ein, wenn Sie an der Verlosung teilnehmen möchten:

Vielen Dank für Ihre Teilnahme!  
Bitte klicken Sie nun auf "Absenden", um die Daten zu übermitteln.

Falls Sie noch Kommentare oder Anregungen haben, können Sie diese hier hinterlassen.

<< Zurück Absenden

Abbildung 15: Letzte Seite der Umfrage mit Verabschiedung und Kommentarfeld

## 9.5 RÜCKLAUFQUOTE

Nach Ablauf des Testzeitraumes von 19 Tagen konnten in den verschiedenen Fragebögen die folgenden Rücklaufquoten erreicht werden:

Kategorie	Anzahl Datensätze	Leere Datensätze	Abbruch	Vollständige Antworten
<b>E-Commerce</b>	33	5	6	22
<b>Entertainment</b>	47	4	2	41
<b>E-Learning</b>	31	5	7	19
<b>Social Software</b>	36	4	2	30
<b>Information</b>	36	1	3	32
<b>Repräsentation</b>	35	6	2	27

Tabelle 25: Rücklaufquote der Umfrage nach Ablauf des Zeitraums von 19 Tagen

Die Anzahl gültiger Antwortsätze pro Kategorie weicht von der Rücklaufquote ab, da bei LimeSurvey auch für aufgerufene Umfragen, die nicht ausgefüllt wurden, ein Datensatz angelegt wird. Die Anzahl leerer Datensätze war allerdings begrenzt.

Die klare Mehrheit an Teilnehmern und vollständig ausgefüllten Fragebögen erreichte die Kategorie Entertainment. Die höchste Zahl unvollständiger Fragebögen konnte in den Bereichen E-Commerce und E-Learning vermerkt werden<sup>41</sup>. Die Gesamtzahl der Teilnehmer belief sich auf

<sup>41</sup> Ein Bogen gilt bei LimeSurvey als unvollständig, wenn ein oder mehrere Items nicht beantwortet wurden.

193 Personen und betrug im Mittel pro Umfrage 31 Teilnehmer. Dies stellt in Anbetracht der Anzahl an Kategorien pro Kategorie eine mittlere, insgesamt jedoch eine akzeptable Rücklaufquote dar. Die Abbruchquote war in der Umfrage relativ gering, mit Ausnahme bei den Kategorien E-Commerce und E-Learning.

## 9.6 FAZIT

Im vorangegangenen Abschnitt konnten die Fragebogenitems, die anhand einschlägiger Usability-Kriterien aus der Literatur entwickelt worden waren, mit Nutzern getestet werden. Hierzu wurden diese mit LimeSurvey erstellt und online zugänglich gemacht. Die Rücklaufquote nach 19 Tagen kann bezüglich der Gesamtzahl an Teilnehmern als akzeptabel, in den einzelnen Kategorien zum Teil jedoch als kritisch betrachtet werden. Dies soll in der in Abschnitt 10 folgenden, statistischen Auswertung der Daten und einer kritischen Diskussion der Ergebnisse weiter erläutert werden.

# 10 STATISTISCHE AUSWERTUNG UND GÜTE DES FRAGEBOGENS

Wie bereits in Abschnitt 5.2 beschrieben, stellen die Hauptgütekriterien der Testtheorie (Objektivität, Reliabilität, Validität) die wichtigsten Qualitätskriterien eines konstruierten Fragebogens dar. Im Folgenden soll anhand einer statistischen Auswertung der Ergebnisdaten überprüft werden, ob die konstruierten Itemsets für den späteren Einsatz tauglich sind. Hierzu sollen zunächst die Maße der zentralen Tendenz und der Dispersion ermittelt werden. Im zweiten Schritt soll eine Analyse des Fragebogens unter den Gesichtspunkten der Haupt- und Nebengütekriterien vorgenommen und die Ergebnisse im Anschluss diskutiert werden.

## 10.1 DESKRIPTIVE STATISTIK ÜBER DIE EINZELNEN ITEMS

Unter der Prämisse, dass eine Likert-Skala statistisch betrachtet als Intervallskala gelten kann<sup>42</sup> sollen im Nachfolgenden die Kennzahlen Antwortverteilung, Mittelwert und Standardabweichung beschrieben und analysiert werden. Hierzu wurden die einzelnen Kategorien einer Häufigkeitsanalyse mit dem Datenauswertungsprogramm SPSS unterzogen und die entsprechenden Kenngrößen ermittelt. In den einzelnen Abschnitten über die Kategorien finden sich beispielhaft Verweise auf die Werte einzelner Items, der komplette Satz an Histogrammen und Tabellen findet sich in Appendix G. Die evaluierte Website findet sich jeweils im Titel, und wird, da sie für die statistische Auswertung keine essentielle Rolle spielt, nur exemplarisch im Text erwähnt.

<sup>42</sup> Zulassung von Differenzbildung, Relationen sowie Aussagen über Unterschiede möglich, vgl. [Bühner2004:70]

Die Kategorie E-Commerce erzielte eine Teilnehmerzahl von insgesamt 28 Personen (vollständige + abgebrochene Umfragen). Zur statistischen Auswertung umgepolt wurden insgesamt vier Items. Alle anderen Items waren positiv formuliert und konnten in ihrer Polung belassen werden<sup>43</sup>. Die Antwortmöglichkeit „keine Angabe/nicht zutreffend“ wurde sowohl im Bereich E-Commerce als auch in allen anderen Kategorien als fehlender Eingabewert („Missing“) behandelt, da ihre Berücksichtigung den Mittelwert und die Standardabweichung verzerren würde.

Der Anteil der Items mit hohen Nonresponse-Quoten war im Bereich E-Commerce deutlich erhöht (15 von 33 Items wiesen Nonresponse-Werte größer oder gleich zwölf auf<sup>44</sup>). Bis auf drei dieser Items<sup>45</sup> ließen sich alle Items mit hohen Nonresponse-Quoten den thematisch übergeordneten Bereichen „Funktionen für registrierte Kunden“ und/oder „Bestellvorgang“ zuordnen. Beispielhafte Items sind hierbei z.B. *„Der Bestellvorgang umfasst zu viele Schritte.“* oder *„Ich weiß, wie ich Angaben zu meinem Kundenkonto ändere.“*. Das lässt sich darauf zurückführen, dass - anders als im Bereich Social Software/Portale – kein Testaccount für die Nutzer angelegt werden konnte<sup>46</sup>, und die Schritte des Bestellvorganges nur nach Anmeldung zu durchlaufen waren. Dies wurde beim Test der Fragebögen nicht berücksichtigt, und ist bezüglich der Auswertung als methodischer Nachteil aufzufassen. Zur Beurteilung der Güte der Items muss in diesem Fall auf die Expertenurteile zurückgegriffen werden. Die Aussagen aus diesem Bereich wiesen nach der Überarbeitung im Durchschnitt hohe Relevanzwerte auf (vgl. Abschnitt 8.7) was bezüglich der Inhaltsvalidität (siehe Abschnitt 10.1.8) ein wichtigeres Kriterium darstellt. Beim Einsatz des tatsächlichen Fragebogens in der Praxis sollte eine statistische Bewertung dieser Items auf jeden Fall nachgeholt werden. Freitextkommentare kamen in der Kategorie E-Commerce an zwei Stellen vor, und bezogen sich auf Funktionen, die abgefragt, jedoch auf der Seite nicht vorhanden waren oder auf Items, die nur als registrierter Nutzer der Seite zu bedienen waren. Die Notwendigkeit eines weiteren Tests mit den Items wird an dieser Stelle nochmals in den Vordergrund gerückt.

Die Seite wurde mit einem Durchschnittsmittelwert von 3,39 bewertet, wobei vor allem die Auffindbarkeit und das Ausreichen der Hilfefunktion (die Seite bietet keine explizit als solche bezeichnete Hilfefunktion an) sowie die Startseite negativ bewertet wurden. Auffälligkeiten in den Antworten ergaben sich bei vier Items hinsichtlich einer zweigipfligen Verteilung<sup>47</sup>. Da diese als möglicher Indikator für eine unterschiedliche Bedeutungsinterpretation bei den Befragten gelten kann [vgl. Bühner 2006:73ff.], wäre als Konsequenz hieraus eine Hilfe oder Erläuterung zu den Items denkbar, um mögliche Unklarheiten (z.B. Begriffe wie „Kategorisierung“ oder „Empfehlung“) bereits im Vorfeld zu klären. Item 30 ist bezüglich seiner Aussage sehr subjektiv, die zweigipflige Verteilung könnte insofern als Geschmacksfrage interpretiert werden. Zu Item 29

<sup>43</sup> Spanne der Polung von den Werten 1 (stimme gar nicht zu) bis 5 (stimme vollkommen zu), sowie -99 (keine Angabe). Umpolung der Itemwerte eins bis fünf.

<sup>44</sup> zwölf oder mehr fehlende Werte bei allen Antworten für ein Item

<sup>45</sup> *„Die Hilfefunktion ist ausreichend für mich.“; „Ich empfinde die Testberichte anderer Kunden als hilfreich.“; „Die Suchfunktion liefert mir nützliche Ergebnisse.“*

<sup>46</sup> Dies erfordert in der Mehrzahl der Fälle die Eingabe von privaten Kundendaten, wie z.B. Angaben zu Adress- und/oder Kontodaten

<sup>47</sup> Da zwei von diesen unter die Gruppe von Items mit hoher Nonresponse fallen, ist eine fundierte statistische Aussage über ihre Güte nur eingeschränkt möglich

lässt sich festhalten, dass die Seite keine konkreten Empfehlungen anbietet, jedoch zu jeder Produktkategorie ausgewählte Produkte auf der Seite anzeigt. Vor allem in einem solchen Fall, bei dem die Deutung des Begriffs „Empfehlung“ unterschiedliche Interpretationen zulässt, sind Kommentare oder Erläuterungen zu den Items unabdingbar.

Insgesamt lässt sich sagen, dass für die Kategorie E-Commerce bedingt durch die begrenzte Teilnehmerzahl und die hohe Nonresponse-Quote bei einigen Items fundierte Schlussfolgerungen nur eingeschränkt möglich ist. Die Antwortverteilungen sind allerdings größtenteils eindeutig. Die Freitextkommentare und die Nonresponse-Quoten lassen den Schluss zu, dass eine weiterführende Analyse im Praxiseinsatz notwendig und wünschenswert ist.

Item	Umgepolte Items
12	„Der Bestellvorgang umfasst zu viele Schritte.“
13	„Ich muss unnötig viele Angaben im Bestellvorgang machen.“
23	„Über die Weitergabe meiner Daten an Dritte mache ich mir auf dieser Seite Sorgen.“
33	„Abbildungen werden auf meinem Bildschirm manchmal nicht angezeigt.“

Tabelle 26: E-Commerce - Umgepolte Items

Item	Auffällige Antwortverteilungen
9	„Die Kategorisierung der Produkte ist übersichtlich.“
20	„Ich weiß, welcher Klick den Kauf auslöst.“
29	„Die Empfehlungen des Systems sind interessant für mich.“
30	„Die Seite wirkt auf Anhieb übersichtlich auf mich.“

Tabelle 27: E-Commerce - Items mit auffälliger Antwortverteilung

### 10.1.2 ENTERTAINMENT: WWW.YOUTUBE.COM

Der Bereich Entertainment wies mit Abstand die größte Teilnehmerzahl auf. Neben der eventuellen Ungleichverteilung des zuvor beschriebenen .php-Skripts auf dem Server wären mögliche Erklärungen hierfür zum einen die Zielgruppe der angesprochenen Umfrageteilnehmer (größtenteils Studierende) und zum anderen die Popularität der Website. Insgesamt konnten 43 Antwortsätze erzielt werden. Zur besseren Auswertung wurde ein Item umgepol. Die Seite wurde mit einem Durchschnittswert von 3,51 bewertet, wobei vor allem die Auffindbarkeit von Funktionen eher negativ auffiel. Eine erhöhte Nonresponse-Quote ergab sich bei vier Items (s.u.), die sich auf die Hilfefunktion oder Anleitungen zur Seite bezogen. Die bewertete Entertainment-Seite bietet keine Tour durch das System an, die Systemhilfe und eine Anleitung befinden sich, eher schwer zu finden, in der unteren horizontalen Navigation. Hierdurch könnte, zusammen mit dem hohen Bekanntheitsgrad und der einschlägigen Hauptfunktion (www.youtube.com wird fast ausschließlich zum Abspielen von Videos verwendet) die Theorie aufgestellt werden, dass die Nutzer die Hilfe entweder nicht benötigen und/oder nicht gesehen haben. Da nur der Durchschnittswert für das Item 17 eine deutlich negative Tendenz aufweist (~2,41), könnte die Vermutung aufgestellt werden, dass die Nutzer die Hilfe (in der Navigation als solche bezeichnet)

eher gefunden haben als die Anleitung (unter dem Navigationspunkt Hilfe als Youtube - Handbuch bezeichnet). Die restliche Anzahl fehlender Angaben pro Items beträgt zwischen null und sieben, was einem aus-reichend zufriedenstellenden Anteil entspricht.

Eine auffällige Ausprägung hinsichtlich einer zweigipfligen Antwortverteilung weist ein Item auf. Eher könnte dies jedoch durch subjektives Empfinden zu erklären sein als durch eine nicht eindeutige Itemformulierung. Ein Item weist eine annähernd gleichverteilte Antwortausprägung auf. Da die evaluierte Website keine Guided Tour aufweist, und das Item zur Gruppe derjenigen mit einer hohen Nonresponse-Quote gehört, sollte dies nicht als Indikator für die Aussagekraft betrachtet werden.

Die Betrachtung der Freitextkommentare zeigt vor allem, dass das Anschreiben bezüglich der Funktionalitäten, die abgefragt, aber nicht vorhanden waren, deutlicher formuliert hätte werden müssen. So war einigen Nutzern nicht klar, dass nicht alle abgefragten Funktionalitäten auch auf der Seite vorhanden waren. Dies ist ein Hinweis dafür, im Handbuch oder der Anleitung zum späteren Programm deutlich zu machen, nur Funktionalitäten der Seite abzufragen, die auch existent sind. Weiterhin wurden an einer Stelle der Hinweis auf die Trennung von Daten und E-Mailadresse sowie der Wunsch nach deutlicherer Hervorhebung der Datenschutzerklärung geäußert. Dies sollte für weitere Fragebogentests eventuell berücksichtigt werden.

Insgesamt kann die Rücklaufquote für die Kategorie Entertainment als zufriedenstellend bezeichnet werden, die Itemanalyse zeigt, dass alle Items ausreichend verständlich und eindeutig formuliert sind. Es lässt sich jedoch festhalten, dass mit dem Ziel, die Itemantworten eindeutiger interpretierbar zu machen, Hinweise an den Items angebracht wären. So ließe beispielsweise eine Anmerkung auf die Platzierung einer Funktionalität (z.B. Hilfe, Anleitung) den Ausschluss des möglichen Grundes „Unauffindbarkeit“ für die Erklärung der Nonresponse-Quote zu.

Item	Umgepolte Items
18	„Auch nach mehrmaliger Benutzung verwirrt mich das System.“

Tabelle 28: Entertainment - Umgepolte Items

Item	Hohe Nonresponse-Quote
14	„Die Hilfefunktion ist ausreichend für mich.“
15	„Die Hilfefunktion ist leicht zu finden.“
16	„Die Guided Tour konnte mir neue Funktionen zeigen.“
17	„Anleitungen und Erklärungen sind leicht zu finden.“

Tabelle 29: Entertainment – Items mit hoher Nonresponse-Quote

Item	Auffällige Antwortverteilung	Verteilung
22	„Die Gestaltung der Seite regt mich dazu an, auf ihr zu surfen (z.B. durch Empfehlungen).“	zweigipflig
16	„Die Guided Tour konnte mir neue Funktionen zeigen.“	gleichverteilt

Tabelle 30: Entertainment – Items mit auffälliger Antwortverteilung



Im Bereich E-Learning war die Teilnehmerzahl mit deutlichem Abstand am geringsten. Erzielt werden konnten 26 gültige Antwortsätze. Dies lässt sich möglicherweise auf die Ungleichverteilung der Teilnehmer auf das Skript zurückführen, zum anderen ist die Sprache Spanisch, die auf der Seite vermittelt werden soll, im Gegensatz zu Englisch möglicherweise nicht für jeden der Teilnehmer von Interesse. Dies sollte bei weiteren Tests mit dem vorliegenden Itemset berücksichtigt werden. Zur statistischen Auswertung wurden insgesamt fünf Items umgepolt. Die Seite wurde mit einem Durchschnittswert von 3,47 bewertet. Der Anteil der Items mit fehlenden Werten lag Bereich E-Learning mit insgesamt sieben Items zwar nicht im optimalen, jedoch immer noch akzeptablen Bereich. Zurückzuführen ist die erhöhte Anzahl fehlender Werte vor allem auf fehlende Funktionalitäten der Seite, was auch aus den Freitextkommentaren hervorgeht. Alle Items bis auf eines<sup>48</sup> fragten Funktionen ab, die das System nicht anbot. Problematisch erweist sich eher die hohe Abbruchquote, so dass von 26 Teilnehmern nur 19 die Umfrage bis zum Ende ausgefüllt hatten. Dies lässt fundierte Schlüsse nur eingeschränkt zu und zeigt den Bedarf nach weiteren Tests im späteren Praxiseinsatz.

Betrachtet man die Antwortverteilungen über die einzelnen Items, so finden sich Auffälligkeiten hinsichtlich einer zweigipfligen Verteilung bei vier Items. Zu überdenken wäre in Konsequenz die Überarbeitung der Formulierung von Item 9 (Verwendung einer Positivformulierung<sup>49</sup>) und 24 (Vereinfachung der Formulierung<sup>50</sup>). Für die Items 1 und 7 wäre eine weitere statistische Analyse nach einem Praxistest mit mehr Teilnehmern wünschenswert, um zu prüfen, ob die auffällige Verteilung auf die begrenzte Anzahl an Personen zurückzuführen ist. Andernfalls wäre bei Item 7 eventuell eine Umformulierung des Items<sup>51</sup> angebracht.

Die Freitextkommentare fielen im Bereich E-Learning eher kritisch aus. Bemängelt wurde zum einen das Fehlen einer konkreten Arbeitsanweisung (ein Hinweis, der in den anderen Kategorien nicht, im Bereich E-Learning jedoch zweimal auftauchte), zum anderen, wie in den Kategorien E-Commerce und Entertainment bereits beschrieben, die Abfrage von Funktionen, die auf der Seite nicht angeboten wurden. Der Entwurf einer konkreten Aufgabe für den Nutzer kann somit vor allem für den Bereich E-Learning als wichtig und notwendig, aber auch für die anderen Kategorien als Option betrachtet werden.

Zusammenfassend lässt sich sagen, dass im Bereich E-Learning nur eine sehr begrenzte Anzahl Teilnehmer zur Verfügung stand, und die evaluierte Seite nicht alle abgefragten Funktionen anbot. Eine aussagekräftige statistische Analyse ist daher nur eingeschränkt möglich. Es lässt sich daraus jedoch nicht zwingenderweise eine Schlussfolgerung auf mangelnde Güte der verwendeten Items ziehen. Trotzdem zeichnen sich bei einigen Items Tendenzen ab, die auf

<sup>48</sup> „Die verwendeten Audio- und Videodateien erleichtern mir das Lernen.“

<sup>49</sup> z.B. „Mir ist jederzeit klar, wie ich an den Anfang einer Lerneinheit zurückkomme.“

<sup>50</sup> z.B. „Ich weiß, wo ich die gewünschten Lerneinheiten finde.“

<sup>51</sup> z.B. „Ich kann die Lerneinheit problemlos wechseln.“

möglichen Überarbeitungsbedarf hinweisen, und die Notwendigkeit einer weiteren Analyse mit mehr Teilnehmern als sehr wünschenswert herausstellen.

Item	Umgepolte Items
6	„Die Aufmachung der Seite lenkt vom Lernstoff ab.“
9	„Manchmal ist mir unklar, wie ich an den Ausgangspunkt einer Lerneinheit zurückkomme.“
16	„Absolvierte Lerninhalte würde ich gerne speichern.“
20	„Ich wünsche mir eine Tour durch das System.“
21	„Texte würde ich mir lieber ausdrucken und auf Papier lesen.“

Tabelle 31: E-Learning – Umgepolte Items

Item	Auffällige Antwortverteilungen
1	„Die Benutzeroberfläche ist übersichtlich.“
7	„Ich kann problemlos zwischen Lerneinheiten navigieren.“
9	„Manchmal ist mir unklar, wie ich an den Ausgangspunkt einer Lerneinheit zurückkomme.“
24	„Ich weiß, welche Lerneinheiten wo zu finden sind.“

Tabelle 32: E-Learning – Auffällige Antwortverteilungen

#### 10.1.4 INFORMATION: WWW.SPIEGEL-ONLINE.DE

Insgesamt konnte in der Kategorie Information eine Rücklaufquote von 32 Antwortsätzen erzielt werden, was einem guten Wert entspricht. Umgepolte wurden drei Items. Der Anteil von Items mit hoher Nonresponse-Quote war mit zwei Items relativ gering, und lässt sich bei einem Item ebenso wie in den anderen Kategorien damit erklären, dass die abgefragte Funktion (Artikelempfehlung) von der Seite nicht angeboten wurde.

Die Seite wurde mit einem Durchschnittswert von 3,58 bewertet, wobei vor allem der Wunsch nach Artikelempfehlungen zum Ausdruck kam. Die Antwortverteilung wies Auffälligkeiten bei zwei Items hinsichtlich einer zweigipfligen Verteilung auf. Anzumerken wäre, dass Item 22 eventuell die Zweiteilung und eine Reduzierung auf jeweils nur ein abgefragtes Merkmal erfordert, da möglicherweise ein empfundener Unterschied zwischen den Konzepten „Sachlichkeit“ und „Neutralität“ bestehen könnte. Die zweigipflige Verteilung in Item 25 wiederum kann man als subjektive Einstellung auffassen, die Prämisse einer einschlägigen Ausrichtung muss hier also nicht verlangt werden. Item 20 wiederum weist im mittleren Bereich eine eher homogene Verteilung mit Tendenz zur Zustimmung auf, dieses Item sollte in kommenden Praxistests einer weiteren Untersuchung unterzogen werden. Sollte sich auch bei höherer Teilnehmerzahl keine Aussagekräftigkeit einstellen, so ist die Entfernung des Items zu überdenken. Insgesamt lässt sich für den Bereich Information eine zufriedenstellende Bilanz ziehen. Die Teilnehmerzahl war mit 35 Teilnehmern und wenig fehlenden Items ausreichend, um

statistisch begründete Aussagen vornehmen zu können. Die Antwortverteilung der Items weist darauf hin, dass die Items (mit einer Ausnahme) eindeutig formuliert und verständlich sind. Für den Bereich Information wurden keine Freitextkommentare geäußert, was ebenfalls ein gutes Zeichen bezüglich der Güte der Items darstellt.

Item	Umgepolte Items
2	„Mir fehlen wichtige Themen, über die die Seite keine Information liefert.“
17	„Ich wünsche mir Empfehlungen zu möglichen interessanten Artikeln.“
23	„Auf der Seite ist zu viel Werbung zu sehen.“

Tabelle 33: Information - Umgepolte Items

Item	Auffällige Antwortverteilung	Verteilung
22	„Die Beiträge der Seite sind sachlich und neutral formuliert.“	zweigipflig
25	„Die Möglichkeit, den Autor eines Textes zu kontaktieren, ist für mich wichtig.“	zweigipflig
20	„Die Seite stellt eine wichtige Informationsquelle für mich dar.“	gleichverteilt

Tabelle 34: Information - Auffällige Antwortverteilung

#### 10.1.5 SOCIAL SOFTWARE: WWW.FACEBOOK.COM

Für die allgemeine Beurteilung einer Website der Kategorie Social Software/Portale ist anzumerken, dass, um den Großteil der Funktionen nutzen zu können, eine Anmeldung auf der Seite erforderlich ist. Zu diesem Zweck wurde ein Account auf der Seite [www.facebook.com](http://www.facebook.com) erstellt, mit dem sich die testenden Personen einloggen und auf der Seite navigieren konnten. Dies wurde mit der Überlegung vorgenommen, die Teilnehmer, die keinen eigenen Account auf der Seite hatten, nicht von der Umfrage auszuschließen. Die Teilnehmer hatten die Wahl, sich mit dem Testaccount oder ihrem eigenen Account auf der Seite einzuloggen, und diese dann im Anschluss zu beurteilen. Die Auswahl eines eigenen Accounts oder eines Testaccounts stellt bezüglich der Evaluierung der Funktionen keinen methodischen Unterschied dar.

Im Bereich Social Software/Portale konnten insgesamt 32 gültige Antwortsätze erzielt werden. Zur statistischen Auswertung wurden zwei der Items umgepolt.

Insgesamt wurde die Seite mit einem Durchschnittswert von 3,18 bewertet, der Anteil von Items mit erhöhter Nonresponse-Quote war auf 4 Items begrenzt. Zwei der Items fragten hierbei Funktionen ab, die die Seite nicht anbietet. Zu einem der beiden anderen Items („Die *Hilfefunktion* konnte meine Fragen klären“) lässt sich sagen, dass eine mögliche Erklärung für den erhöhten Anteil fehlender Werte der Bekanntheitsgrad der Website unter der Teilnehmergruppe sein könnten, so dass viele der teilnehmenden Testpersonen keine Hilfestellung zur Navigation auf der Seite benötigten. Weiterhin wird die Hilfefunktion auf der Seite [www.facebook.com](http://www.facebook.com) relativ weit unten in der rechten Ecke angeboten, so dass eine weitere mögliche Erklärung wäre, dass die

Hilfefunktion durch die Nutzer nicht gefunden wurde. Das vierte Item mit erhöhter Nonresponse-Quote („Die Kommentierung oder Verschlagwortung von Beiträgen ist einfach“) lässt, wie bereits in vorangehenden Abschnitten beschrieben, die Vermutung zu, dass die alleinige Formulierung des Items für die Testperson nicht vollständig klären konnte, was gemeint ist. Die evaluierte Seite bietet die Funktion der Beitragskommentierung auf der Startseite an, es muss hier also vermutet werden, dass den teilnehmenden Personen durch die Fragestellung nicht klar war, welche Funktionalität der Seite mit der Itemformulierung gemeint war. Hier ist auf jeden Fall die Notwendigkeit für einen Hinweis oder einer genaueren Erläuterung zu sehen. Bezüglich der Antwortverteilung zeigte sich bei fünf Items die Tendenz zu einer zweigipfligen Verteilung. Hier wäre durch weitere Tests in der Praxis und einer größeren Teilnehmerzahl zu prüfen, welche Ergebnisse sich abzeichnen.

Die Freitextkommentare der Kategorie Social Software bezogen sich, wie in den vorangegangenen Abschnitten bereits beschrieben, sämtlich auf die Nichtexistenz von abgefragten Funktionalitäten, Hinweise auf Verständnisprobleme wurden nicht abgegeben.

Zusammenfassend lässt sich sagen, dass sowohl die Rücklaufquote als auch die Nonresponse-Rate in der Kategorie Portale/Social Software zufriedenstellend war. Festzuhalten ist die Notwendigkeit nach einer genaueren Erläuterung zu Items, deren Begrifflichkeit auf Anhieb offenbar nicht eindeutig ist (bspw. der Begriff Verschlagwortung). Interessant wäre weiterhin eine weitere Analyse nach der Durchführung weiterer Praxistests, um die Items, die momentan eine zweigipflige Antwortverteilung aufweisen, genauer analysieren und gegebenenfalls überarbeiten zu können.

Item	Umgepolte Items
14	„Ich würde das Layout der Seite gerne meinen Bedürfnissen anpassen.“
15	„Ich würde mein Nutzerprofil gern ausführlicher beschreiben.“

Tabelle 35: Social Software – Umgepolte Items

Item	Auffällige Antwortverteilungen
1	„Die Benutzeroberfläche ist übersichtlich.“
3	„Die Funktionen zur Kommunikation mit anderen Nutzern sind leicht zu bedienen.“
5	„Die Funktionen der Seite sind auf Anhieb ersichtlich.“
7	„Die Hilfefunktion ist jederzeit auffindbar.“
8	„Die Möglichkeit, das Nutzerprofil zu ändern, ist gut sichtbar platziert.“

Tabelle 36: Social Software – Auffällige Antwortverteilungen

#### 10.1.6 REPRÄSENTATION: WWW.BHPEG.DE

In der Kategorie Repräsentation/Homepages konnten insgesamt 29 gültige Antwortsätze erzielt werden. Zur statistischen Analyse umgepolzt wurden drei der Items. Die Website, der Internetauftritt einer Projektentwicklungs-GmbH, wurde im Durchschnittswert mit 3,31 bewertet. Die Anzahl von Items mit erhöhtem Anteil an fehlenden Angaben (größer gleich 12) belief sich auf insgesamt drei. Die Items mit erhöhtem Wert fehlender Angaben fragten allesamt nach nichtexistierenden Funktionalitäten: Zwei der Items bezogen sich auf die Suchfunktion, die auf der Seite nicht vorhanden ist. Die anderen zwei bezogen sich auf Multimedia-Dateien, die auf der Seite ebenfalls nicht eingesetzt werden. Die fehlenden Angaben können also als Reaktion auf die Nichtexistenz der Funktionalitäten auf der Seite betrachtet werden, nicht als Hinweis auf Unverständlichkeit der Itemformulierung oder Bedeutungsunklarheiten.

Bei der Antwortverteilung ergaben sich bis auf ein Item, welches eine nichtexistente Funktion abfragte, keine Besonderheiten oder Tendenzen zu einer zweigipfligen Verteilung. Die Items können also als schlüssig und eindeutig formuliert betrachtet werden.

Die Freitextkommentare der Kategorie Repräsentation bezogen sich fast ausschließlich auf die Seite selber, weniger auf die Items und den Fragebogen. An zwei Stellen kamen die bereits aus den anderen Kategorien bekannten Hinweise auf die Nichtexistenz abgefragter Funktionalitäten.

Zusammenfassend lässt sich sagen, dass die Rücklaufquote trotz des geringen Anteils von Items mit erhöhter Nonresponse-Quote mit 29 Antwortsätzen noch zu gering ist, um aussagekräftige und statistisch begründete Aussagen treffen zu können. Es lässt sich jedoch festhalten, dass die bisherigen Tendenzen bei der Antwortverteilung und den Nonresponse-Quoten zufriedenstellend sind. Dies sollte jedoch in weiteren Praxistests mit einer größeren Teilnehmerzahl getestet werden.

Item	Umgepolzte Items
10	„Auf der Seite ist zu viel Werbung zu sehen.“
15	„Mein Computer benötigt lange, um Bilder und Seitenelemente zu laden.“
16	„Dieselben Informationen werden unnötig oft geliefert.“

Tabelle 37: Repräsentation – Umgepolzte Items

Item	Auffällige Antwortverteilungen	Verteilung
7	„Die Suchfunktion liefert nützliche Ergebnisse.“	zweigipflig

Tabelle 38: Repräsentation – Auffällige Antwortverteilungen

## 10.2 ZUSAMMENFASSUNG

Insgesamt ist festzuhalten, dass die Analyse der deskriptiven Statistik zufriedenstellende Ergebnisse für die Itemsets lieferte, insoweit dies möglich war. Hohe Nonresponse-Quoten oder auffällige Verteilungen ließen sich größtenteils durch die Nichtexistenz abgefragter Funktionen

auf der evaluierten Website erklären. Allerdings zeichnete sich die Notwendigkeit nach Hinweisen oder Erklärungen zu einzelnen Items bei der Auswertung ab. Nachfolgend soll nun die Analyse der Ergebnisse nach den Kriterien der klassischen Testtheorie erfolgen, um die Güte der erstellten Items zu erforschen.

### 10.3 DISKUSSION DER FRAGEBÖGEN UNTER DEN ALLGEMEINEN GÜTEKRITERIEN

Im Vergleich zur Überprüfung von Reliabilität und Validität klassischer Tests und Fragebögen nach der allgemeinen Testtheorie muss bei den vorliegenden Itemsets angemerkt werden, dass sich diese inhaltlich deutlich von psychometrischen Tests zur Messung bestimmter psychologischer Konstrukte (Merkmalsausprägungen) unterscheiden. Ziel der späteren Fragebögen ist es weder, ein Persönlichkeitsmerkmal beim Probanden zu messen, noch eine Hypothese zu verifizieren oder zu falsifizieren. Sinn der Evaluation soll sein, nach der Durchführung Signale und Hinweise auf mögliche Schwachstellen einer Website zu erhalten. Zu diesem Zweck wurden Items konstruiert, die gemäß den Hinweisen von Literatur und Experten einzelne Funktionalitäten einer Website abfragen, und somit inhaltlich weitgehend unabhängig voneinander sind.

Als Beispiel sei eine E-Commerce-Website angeführt, deren Suchfunktion prominent platziert ist und zufriedenstellende Ergebnisse liefert, deren Warenkorb jedoch für den Nutzer nicht zu finden ist. Beide Items weisen keine inhaltliche Korrelation auf. Ähnliches lässt sich für zwei Items desselben „Themengebiets“ aussagen, da die Platzierung einer Funktionalität nicht automatisch die Eigenschaften ihrer Bedienung impliziert<sup>52</sup>. Die Items sind also inhaltlich stark heterogen und weitgehend unabhängig voneinander, und müssen in einer entsprechenden Analyse der Testergebnisse auch so behandelt werden. Diese Eigenschaft der konstruierten Itemsets bedingt es, den Sinn einer klassischen Reliabilitäts- und Validitätsanalyse zu überdenken. Dies soll anhand des nachfolgenden Abschnittes verdeutlicht werden.

Die Objektivität von Fragebogenverfahren im Internet wurde bereits in Abschnitt 5.3.2 erörtert. Bedingt durch den standardisierten Ablauf, die räumliche Distanz zwischen dem Testleiter und dem Testenden sowie die automatische maschinelle Auswertung mit festgelegter Polung bei Likert-Skalen sind alle drei Objektivitätskriterien<sup>53</sup> erfüllt. Die Ergebnisse bedürfen an dieser Stelle keiner weiteren Analyse.

#### 10.3.1 RELIABILITÄT

Die Reliabilität eines Fragebogens ist, wie in Abschnitt 5.2 beschrieben, der Grad der Reproduzierbarkeit der Messergebnisse und die Frage nach der Korrelation mehrerer Ergebnisse desselben Tests. Es ist festzuhalten, dass zum einen aus zeitlichen Gründen verschiedene

<sup>52</sup> Beispiel: Warenkorb leicht zu finden, Möglichkeit, Produkte zu entfernen jedoch kaum ersichtlich.

<sup>53</sup> Durchführungsobjektivität, Auswertungsobjektivität, Interpretationsobjektivität

Verfahrensweisen, wie Testhalbierungsreliabilität<sup>54</sup> nicht durchführbar waren oder sich zum anderen, wie Retest-Reliabilität<sup>55</sup>, Paralleltestreliabilität<sup>56</sup> oder interne Konsistenz in der Methode als nicht sinnvoll erweisen.

Zum einen soll eine Benutzerbefragung zur Usability-Evaluation dazu dienen, Fehler in der Gebrauchstauglichkeit, die vor allem für Nutzer ohne Vorwissen hinderlich sind, aufzudecken. Insofern würde eine erneute Durchführung des Usability-Tests mit den gleichen Teilnehmern (Retest) die Ergebnisse beeinträchtigen, da die Nutzer bereits im Umgang mit der Website vertraut wären. Ob im Rahmen der vorliegenden Arbeit für einen derartigen Test erneut genug Nutzer zur Verfügung stünden, ist ebenfalls fraglich.

Für die Erhebung der Halbierungsreliabilität war die Rücklaufquote pro Kategorie im Mittel zu gering. Für eine exploratorische Datenanalyse der Ergebnisse im späteren Praxiseinsatz ist eine Anzahl von rund 30 Teilnehmern durchaus geeignet, erste Indikatoren auf Usability-Probleme aufzuzeigen. Für eine statistisch fundierte Analyse der Daten hinsichtlich von Verteilungswerten ist dieser Stichprobenumfang jedoch ungeeignet. Eine versuchsweise Durchführung mit SPSS für die Rücklaufdaten der Kategorie mit den meisten Antworten (Entertainment) ergab keine aussagekräftigen Ergebnisse<sup>57</sup>.

Ebenso wenig ist es sinnvoll, mit inhaltlich größtenteils voneinander unabhängigen Items eine Paralleltestreliabilität zu messen. Die Generierung zweier Itemmengen, die inhaltlich exakt vergleichbar sind, ist im Allgemeinen als schwierig einzustufen [vgl. auch Bühner 2006:119]. Bei Aufteilung des Itempools wäre bei den vorliegenden Items keine Messung des gleichen Sachverhaltes durch die beiden Itemgruppen gewährleistet.

Die Messung des internen Konsistenzkoeffizienten soll, da sie in der Mehrzahl konstruierter Usability-Fragebögen ermittelt wird, ebenfalls angesprochen werden. Bühner (2006) verweist darauf, dass die Messung des Korrelationskoeffizienten mit inhaltlich heterogenen Items nicht sinnvoll ist [ebd.]. Eine probeweise Durchführung der Reliabilitätsberechnung bestätigt diese Aussage.

Nachfolgend die Werte für die 6 Kategorien:

Kategorie	Fälle	Davon gültig <sup>58</sup>	Items	Korrelationskoeffizient
<b>Entertainment</b>	43	34	24	0,712
<b>E-Commerce</b>	28	21	33	0,904
<b>Information</b>	35	32	25	0,574
<b>E-Learning</b>	26	17	28	0,835

<sup>54</sup> Korrelation/Grad der Übereinstimmung der Ergebnisse zweier Hälften desselben Tests bei Teilung

<sup>55</sup> Korrelation der Testleistungen bei mehrfacher Durchführung des Tests unter denselben Bedingungen

<sup>56</sup> Korrelation zwischen zwei Tests bei Messung desselben Merkmals mittels unterschiedlicher Items

<sup>57</sup> Korrelationswerte <0,22, s. Appendix I

<sup>58</sup> Da das Programm SPSS Items mit hoher Anzahl fehlender Werte systematisch ausschließt, weicht die Zahl der Items, über die die Analyse durchgeführt werden kann, zum einen für alle Kategorien von der Zahl der tatsächlichen Items ab

<b>Social Software</b>	32	30	23	0,898
<b>Repräsentation</b>	29	25	16	0,659

**Tabelle 39: Korrelationskoeffizienten für die sechs Kategorien**

Die Korrelationskoeffizienten variieren stark zwischen den einzelnen Kategorien, die höchsten Werte sind hierbei im Bereich E-Commerce und Social Software zu vermerken. In den Kategorien Information und Repräsentation, finden sich die geringsten Werte. Inhaltlich gesehen lässt sich dieses Ergebnis nicht als Güte der einzelnen Items interpretieren. Vielmehr kann, durch die starke inhaltliche Heterogenität der Items, die Hypothese aufgestellt werden, dass ein hoher Reliabilitätskoeffizient einer Kategorie darauf hinweist, dass die von den Testern wahrgenommene Usability über alle abgefragten Funktionen annähernd ähnlich ist. Dies impliziert allerdings keine Wertung, sondern sagt lediglich dass eine Seite durchgehend als gleich gut, oder gleich schlecht wahrgenommen wird. Analog lässt sich die Hypothese auf einen niedrigen Korrelationskoeffizienten anwenden: Je geringer der Wert, desto heterogener die Usability der Seite, dies bedeutet, die Gebrauchstauglichkeit einzelner Funktionen und Elemente weicht stark voneinander ab. Zur Interpretation des Korrelationskoeffizienten in einem solchen Kontext müssten zusätzlich die Mittelwerte der kategoriespezifischen Items herangezogen werden, die eine Aussage über die abgegebene Bewertung der Nutzer ermöglichen.

Zusammengefasst lässt sich also festhalten, dass die Bewertung der vorliegenden Items unter dem Gesichtspunkt Reliabilität mit einigen Schwierigkeiten verbunden ist. Die bestehenden Methoden erfordern entweder eine größere Stichprobe als erreicht wurde oder sind im vorliegenden Kontext nicht anzuwenden. Da sich nach einer eingehenden Diskussion der Methoden jedoch der Eindruck herauskristallisiert, dass eine Analyse, die berechnet, wie hoch die Ergebnisse von Teilen eines Tests oder einzelner Items miteinander korrelieren, im vorliegenden Kontext wenig Sinn macht, ist dies in der vorliegenden Arbeit akzeptabel.

### 10.3.2 VALIDITÄT

Validität stellt das Hauptgütekriterium eines konstruierten Fragebogens dar. Unter Validität wird, wie in Abschnitt 5.2 beschrieben, der Grad der Messgenauigkeit verstanden, also inwieweit das zu erfassende Konstrukt wirklich gemessen wird.

Da im Vorfeld intensive Literaturrecherche zur Usability-Forschung in den einzelnen Website-Kategorien betrieben wurde, kann davon ausgegangen werden, dass der zu erfassende Sachverhalt (kategoriespezifische Usability) ausreichend abgedeckt ist. Die jeweiligen Itempools können entsprechend als inhaltsvalide betrachtet werden. Auch die Revision der erstellten Items durch Experten und Studenten trägt zusätzlich zur Inhaltsvalidität bei.

Da die Kriteriumsvalidität, wie bereits dargelegt, die Korrelation der Testergebnisse mit einem externen Faktor oder einer Messzahl darstellt, ist dies im Kontext von Websites und mit den



vorliegenden Testergebnissen eher schwierig zu beurteilen. Denkbar wäre hier ein einzelfallspezifischer Abgleich von den Testergebnissen ausgewählter Items mit beispielsweise Besucherzahlen und -frequenz, Dauer des Besuchs auf der jeweiligen Website oder Umsatzzahlen. Unter Voraussetzung einer entsprechenden Rücklaufquote lassen sich so mit hoher Wahrscheinlichkeit aussagekräftige Ergebnisse erzielen. Da dies jedoch größtenteils Daten sind, die den Betreibern der Website vorliegen, sollte dies im Rahmen des späteren konkreten Praxiseinsatzes unter Einbeziehung des Betreibers geschehen.

Zur Erhebung der Konstruktvalidität, also inwieweit die erhobenen Testergebnisse mit anderen Testergebnissen zum gleichen Sachverhalt korrelieren, könnte in der Usability-Evaluation beispielsweise ein Vergleich der Befragungsergebnisse mit den Resultaten einer Experten-Evaluation durchgeführt werden. Denkbar ist dies beispielsweise im Rahmen einer integrierten Zusammenarbeit unter Anwendung des Expert-Review-Tools, welches ebenfalls auf der Plattform [www.usability-toolkit.de](http://www.usability-toolkit.de) angeboten wird. Auch hier ist eine entsprechend hohe Rücklaufquote sowie ein für den konkreten Anwendungsfall erstellter Fragebogen, der nur tatsächlich existierende Funktionen und Elemente einer Website abfragt, Grundvoraussetzung für aussagekräftige Ergebnisse.

---

### 10.3.3 NEBENGÜTEKRITERIEN

Die vier Nebengütekriterien für einen Test sind Normierung, Vergleichbarkeit, Ökonomie und Nützlichkeit. Überprüft man die bestehenden Itemsets diesbezüglich, so steht fest, dass die Kriterien Ökonomie und Nützlichkeit voll erfüllt werden. Die ökonomischen Vorteile von Online-Fragebogenverfahren wurden bereits in Abschnitt 5.3.1 erörtert. Hinsichtlich des Kriteriums der Nützlichkeit ist zu sagen, dass bisher kaum Ansätze zur kategoriespezifischen Beurteilung der Usability von Websites existieren. Insofern können die erstellten Items als vollständig nützlich erachtet werden, was allerdings bedeutet, dass für die einzelnen Kategorien keine Referenzwerte zur Einordnung der ermittelten Testergebnisse vorliegen. Dennoch existieren zahlreiche Fragebögen zur allgemeinen Erhebung der Usability von Websites, so dass die Kriterien Normierung und Vergleichbarkeit zumindest im Rahmen späterer Praxiseinsätze zu erfüllen sind.

## 10.4 FAZIT

Im vorangegangenen Abschnitt konnte die durchgeführte statistische Auswertung der Testergebnisse zeigen, dass die Items in den Kategorien Entertainment, Social Software, Information und Repräsentation zufriedenstellende, wenngleich nicht optimale Rücklaufquoten erzielen konnten.

Die wichtigsten Gütekriterien konnten in den vorliegenden Itemsets erfüllt werden. Es lässt bestätigt sich, dass die ermittelten Fragebogenitems vollständig, verständlich, objektiv und valide

sind. Auch die Nebengütekriterien konnten zumindest teilweise erfüllt werden. Die Überprüfung der Items auf ihre Güte ist somit abgeschlossen. Zum Einsatz der Items in der Praxis soll im letzten Abschnitt ein prototypischer Entwurf für die spätere Anwendung folgen, die die Grundlage für die Umsetzung des Programms bilden soll.

## 11 PROTOTYPISCHE ERSTELLUNG DES FRAGEBOGENS

Nachdem in den vorangegangenen Abschnitten die kategoriespezifischen Items aus der Literatur entwickelt wurden, soll nun die prototypische Entwicklung des Fragebogengenerators erfolgen. Mit Anforderungsanalyse, Konzeption und prototypischem Design soll ein Grundkonzept für eine spätere Implementierung des Fragebogengenerators geschaffen werden.

### 11.1 ANFORDERUNGSANALYSE

Die Anforderungsanalyse soll dazu dienen, fachliche und technische Anforderungen an ein System festzulegen sowie Nutzerwünsche mit einzubeziehen und zu berücksichtigen [s. Starke 2008:29, Kleuker 2009:51]. Bei dem zu erstellenden Programm wird es sich gemäß der Idee des BUI-Projekts um eine Anwendung handeln, die dem Benutzer auf der BUI-Internetplattform [www.usability-toolkit.de](http://www.usability-toolkit.de) zur Verfügung gestellt wird. Den Benutzern, bei denen es sich, wie in Abschnitt 1.1 beschrieben, um kleine und mittelständische Unternehmen ohne Usability-Vorkenntnisse handelt, soll die Möglichkeit geboten werden, einen Fragebogen zur Erhebung der Gebrauchstauglichkeit ihrer Website zu erstellen und Umfragedaten zu erheben. Dem Benutzer werden neben selektierbaren, kategoriespezifischen Fragebogenitems ebenfalls zusätzliche Items angeboten, die dazu dienen, kategoriefremde<sup>59</sup> Eigenschaften oder Funktionalitäten der jeweiligen Website abzufragen. Das Tool stellt weiterhin einen Standardfragebogen mit den wichtigsten Items für jede Kategorie bereit, zusammen mit einer umfassenden Hilfestellung. Ziel des erstellten Fragebogens ist es, dem Nutzer auf möglichst einfache und intuitive Art und Weise zu ermöglichen, eventuelle Schwachstellen und Problemfelder seiner Website aufzudecken und möglichen Handlungsbedarf zu erkennen. Für diesen Zweck wird das Tool zusätzlich die Möglichkeit der Datenauswertung und -visualisierung anbieten.

Hierzu soll im ersten Schritt eine Anforderungsanalyse durchgeführt und der Ablauf des geplanten Programmes in der Konzeption anhand eines Sequenzdiagrammes visualisiert werden. Im Anschluss erfolgt die Erstellung eines Prototyps erfolgen, der, ohne mit weitreichenden Funktionen versehen worden zu sein, die Interaktion des späteren Benutzers mit der Benutzeroberfläche simulieren kann.

---

<sup>59</sup> Gemeint sind hierbei Items, deren vorherrschender inhaltlicher Schwerpunkt in einer anderen Kategorie liegt, denkbar wäre hier z.B. die Bewertung der Kommunikationsfunktionen zwischen Nutzern auf einer E-Commerce-Plattform. Der inhaltliche Schwerpunkt läge im Rahmen dieser Arbeit in der Kategorie „Social Software“

## 11.2 ARCHITEKTUR

Das zu erstellende Programm soll als Webanwendung realisiert werden. Bei einer Webanwendung handelt es sich um ein Programm, das auf einem Webserver gehostet wird und mit dem der Nutzer über einen Webbrowser interagiert. Die Erstellung der grafischen Benutzeroberfläche erfolgt über serverseitig generierten Quellcode (bspw. HTML, XHTML, PHP) aufgrund von HTTP-Requests des Clients nach Start der Webanwendung und während der Interaktion mit dem Benutzer. So entstandene Daten können serverseitig über SQL-Befehle in einer Datenbank gespeichert und ausgelesen werden, nutzerspezifische Daten können als HTTP-Cookies clientseitig gespeichert werden [Kannengiesser 2007:81ff.].

Der zu erstellende Fragebogengenerator wird entsprechend auf dem Webserver des BUI-Projekts gehostet. Sowohl die vom Nutzer generierten Daten aus Fragebogengenerierung und -bearbeitung als auch die aus der Durchführung der Usability-Evaluation (mit dem erstellten Fragebogen) entstandenen Daten werden serverseitig in einer Datenbank abgelegt. Abbildung 16 zeigt die Client-Server-Architektur der zu erstellenden Anwendung.

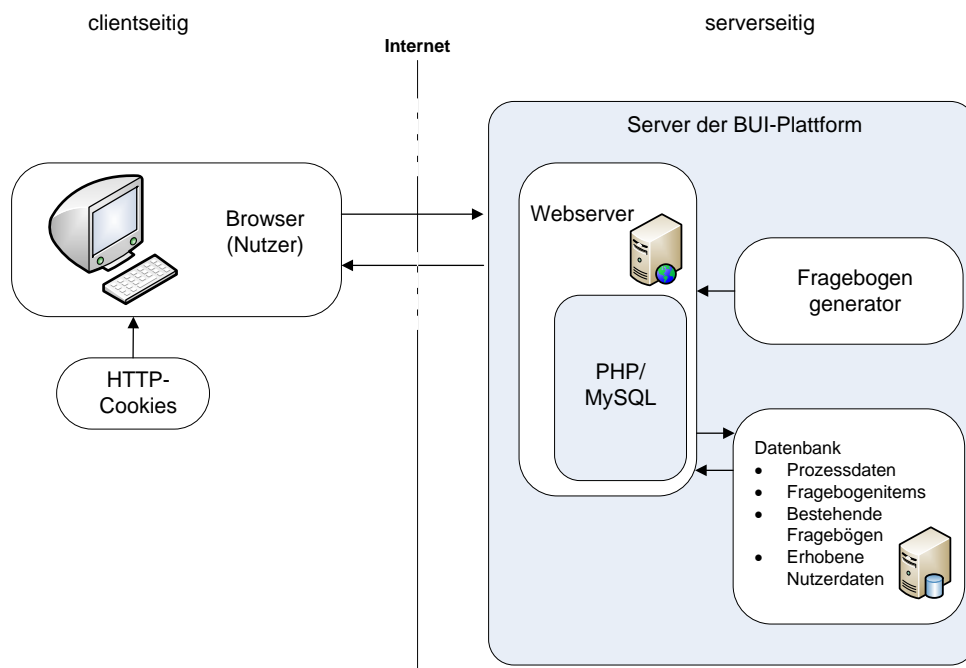


Abbildung 16. Client-Server Architektur des Prototypen

### 11.3 PROGRAMMFUNKTIONEN

Um die in Abschnitt 10.1 beschriebenen Anforderungen zu erfüllen und dem Benutzer ein möglichst einfaches und intuitives Arbeiten mit der Anwendung zu ermöglichen, wird die Anwendung folgende Funktionen beinhalten:

#### *Login/Logout*

Direkt nach dem Start des Programms loggt der Nutzer sich mit Benutzernamen und Passwort ein. Der Login dient dazu, dem Benutzer neben einer Sicherheitsfunktion die Möglichkeit bereitzustellen, bestehende (von ihm erstellte) Bögen zu bearbeiten oder erhobene Daten auszuwerten.

#### *Registrierung*

Ist der Benutzer noch nicht im System angelegt, besteht die Möglichkeit, sich mit Benutzername und Passwort zu registrieren. Denkbar ist entweder eine Anmeldung mit E-Mailadresse, Benutzernamen und Passwort, so dass die Registrierung nach Zusendung einer E-Mail noch bestätigt werden muss, oder eine direkte Registrierung mit Benutzernamen und Passwort.

#### *Erstellung eines neuen Fragebogens*

Nachdem der Benutzer im System eingeloggt ist, gibt es zwei Entscheidungsalternativen bezüglich der auszuführenden Tätigkeit („Neuen Bogen erstellen“ oder „Bogen bearbeiten“). Wird die Option „Neuen Bogen erstellen“ gewählt, wird dem Benutzer zunächst die Möglichkeit geboten, eine der sechs Hauptkategorien, der seine Seite zuzuordnen ist, auszuwählen.

#### *Wahl eines (kategoriespezifischen) Hauptitemsatzes*

Zur Wahl der Hauptkategorie wird dem Benutzer eine Auswahl der sechs möglichen Kategorien präsentiert. An dieser Stelle existieren aktive Systemhinweise (Erklärungen zu den Kategorien und Angabe von Beispielwebsites) in Form von Tooltips. Durch Auswahl einer Kategorie wird der Itemsatz der Hauptkategorie – grafisch als Baumstruktur realisiert – in die GUI geladen. Dieser umfasst alle in einer Kategorie befindlichen Items, welche sich wiederum thematisch-funktional in sogenannte Itempakete untergliedern. Diese Art von Gliederung soll eine intuitive Orientierung des Nutzers gewährleisten, während eine Anordnung der Items unter den Gestaltungsgrundsätzen nach ISO-9241 vermutlich eher zu Verwirrung beim Benutzer führen würde. So enthält beispielsweise das Itempaket „Warenkorb“ alle Items, die inhaltlich der Kategorie Warenkorb zuzuordnen sind. Jeder Kindknoten der Baumstruktur stellt entsprechend ein Itempaket dar. Da die Kindknoten beim Laden der Hauptkategorie zunächst geschlossen sind, soll dies dazu dienen, dem Nutzer auf den ersten Blick einen sinnvollen Überblick zu verschaffen, an welcher Stelle ein Item zu finden ist.

### *Selektion/Deselektion einzelner Items oder kompletter Itempakete*

Im Anschluss kann der Nutzer einzelne Items oder vollständige Itempakete auswählen, die zusammen den späteren Fragebogen ergeben. Die Items des präsentierten Itemsatzes sollen möglichst leicht an- und abzuwählen sein. Realisiert wurde dies an dieser Stelle mit Hilfe von Checkboxen. Alternativ wäre eine farbliche Markierung nach Anklicken des Items oder eine Schriftschnittänderung (fett/kursiv) denkbar.

Das Programm gewährleistet, dass dem Nutzer zu jedem Zeitpunkt visualisiert wird, wie viele Items insgesamt ausgewählt sind, und gibt bei einer Überschreitung der Anzahl von 20 Items einen Warnhinweis bezüglich der Fragebogenlänge aus. Dies wurde durch die Anzeige der Itemzahl zum einen in der Navigation (s. Abschnitt 10.3) und zum anderen direkt in der Baumstruktur umgesetzt. Der Hinweis auf die Fragebogenlänge erfolgte mit einer Visualisierung in horizontaler Ampelform. Die einzelnen Items werden innerhalb der Pakete absteigend nach Relevanz angeordnet, so dass dem Nutzer wichtige Items an oberster Stelle präsentiert werden. Zusätzlich kann an dieser Stelle die Visualisierung der Itemrelevanz durch Icons angedacht werden. Wichtig ist hierbei der Rückgriff auf bereits bestehende mentale Modelle beim Benutzer, um die Intuitivität der Wahrnehmung der Icons zu gewährleisten. Denkbar wäre beispielsweise eine Realisierung mithilfe von Sternen bestimmter Anzahl (je mehr Sterne, desto relevanter das jeweilige Item).

Angewählte Items werden in der Fragebogenvorschau in der Reihenfolge der vorherigen Pakete präsentiert. Die Reihenfolge kann lediglich in der Vorschau verändert werden. Die Veränderung der Itemreihenfolge bereits während der Ansicht „Hauptitemsatz“ ist nicht sinnvoll, da die thematisch-funktionale Gliederung in unterschiedliche Itempakete eine grundlegende Orientierungsfunktion für den Nutzer darstellt. Im Anschluss kann eine erneute Änderung des Bogens, die Speicherung und/oder der Export des bestehenden Bogens erfolgen.

### *Wahl von zusätzlichen (kategoriefremden) Itempaketen*

Dem Benutzer wird jederzeit die Möglichkeit geboten, auch kategoriefremde Itemsätze anzuwählen, um Seitenfunktionalitäten evaluieren zu können, die von der eigentlichen Kategorie abweichen, aber dennoch angeboten werden (denkbar wäre an dieser Stelle beispielsweise ein Diskussionsforum für Kunden in einem Online-Shop). Diese werden als Reiter in der oberen horizontalen Navigation angeboten. Hierbei wurde darauf geachtet, der Übersichtlichkeit halber nur einen Reiter „Zusatzkategorien“ anzubieten, innerhalb dessen dann jeweils die Kategorien ausgewählt werden können (siehe Abschnitt 11.7.5). Die Strukturierung der Items der Nebenkategorie erfolgt ebenfalls als Baumstruktur. Die Selektion und Deselektion erfolgt analog zu den Items der Hauptkategorie. Auch hier ist die Anzahl der Items jederzeit in der Navigation und in der Baumstruktur sichtbar.

### *Daten und Einstellungen*

Der Nutzer hat die Möglichkeit, für den zu erstellenden Fragebogen verschiedene Einstellungen vorzunehmen und Daten einzugeben. Unter anderem werden Textfelder für Titel, Begrüßung, Verabschiedung und Datenschutz angeboten. Der Nutzer kann festlegen, ob es sich um eine öffentlich zugängliche oder für einen bestimmten Teilnehmerkreis erstellte Umfrage handeln soll. Hierbei besteht die Möglichkeit, entsprechende Zugangsschlüssel zu generieren. Die Fragen können als „obligatorisch“ oder als „freiwillige Eingabe“ gekennzeichnet werden. Auch die Veröffentlichung des Fragebogens kann hier vorgenommen werden.

### *Bearbeitung eines bestehenden Fragebogens*

Wählt der Nutzer die Option „Bestehenden Bogen aufrufen“, so wird er zu einem Übersichtsmenü weitergeleitet, in dem die Möglichkeit zur Bearbeitung, Löschung, Export oder Auswertung eines bestehenden Bogens besteht. Die Bearbeitung bestehender Bögen verläuft funktional analog zur Erstellung eines neuen Bogens unter Auslassung der Interaktiven Oberfläche „Hauptkategorie auswählen“. Der angewählte Bogen (die Visualisierung der bestehenden Bögen ist als Dropdown-Liste realisiert) wird direkt in die Benutzeroberfläche geladen und steht dort zur Bearbeitung zur Verfügung. Items können entsprechend selektiert und deselektiert werden, Speicherung und Export ist entweder unter derselben oder einer alternativen Bezeichnung möglich.

### *Löschen eines bestehenden Fragebogens*

Unter dem der Menüpunkt „Bearbeitung eines bestehenden Bogens“ ist es möglich, in der anfänglichen Benutzeroberfläche die Option „Bestehenden Bogen löschen“ anzuwählen. Ein bestehender Bogen kann so endgültig gelöscht werden. Diese Möglichkeit wurde mit einer Sicherheitsabfrage und dem Hinweis, dass sämtliche Daten verloren gehen, versehen. Als Option ist an dieser Stelle eine Trennung zwischen alleinigem Löschen des Fragebogens und zusätzlichem Löschen von zuvor erhobenen Daten möglich. Eine weitere Abfrage („zusätzlich auch alle erhobenen Daten löschen?“) integriert die entsprechende Möglichkeit.

### *Vorschau des erstellten oder bearbeiteten Bogens*

Damit sich der Benutzer einen Überblick über Länge und Aussehen des fertigen Fragebogens verschaffen kann, beinhaltet das Programm die Funktion „Vorschau“. Hier ist auch die Änderung der Reihenfolge der einzelnen Fragebogenitems möglich. Nach der Vorschau soll ein direktes Speichern, der Export, oder eine weitere Bearbeitung des Bogens ermöglicht werden.

### *Speichern des erstellten oder bearbeiteten Bogens*

Nach dem Ende der Fragebogenerstellung oder bei Unterbrechung der Bearbeitung des Bogens kann der Bogen gespeichert werden, um ihn später erneut aufrufen zu können. Von der Speicherfunktion ist die Exportfunktion für den fertigen Fragebogen zu unterscheiden.

### *Export des Bogens*

Zusätzlich zur Speicherung auf dem Datenbankserver der BUI-Plattform ist ein Export des zusammengestellten Fragebogens möglich. Dies stellt insofern eine sinnvolle Ergänzung dar, als dass der Nutzer eine eigene Kopie des von ihm zusammengestellten Itemsatzes erhält, die auch bei versehentlichem Löschen, eventuellen Datenverlusten auf dem Webserver oder bei zukünftigem Rückgriff auf alte Fragebögen zur Verfügung steht. Der Export des Bogens ist als Tabelle (.csv, .xl\*, .xml-Format) oder auch als Textdatei (.doc, .txt) möglich.

### *Veröffentlichung der Benutzerbefragung*

Sobald die Bearbeitung des Bogens abgeschlossen ist, kann die Benutzerbefragung im Menüpunkt „Daten und Einstellungen“ veröffentlicht werden. Hierzu wird dem fertigen Fragebogen vom System eine http – Adresse zugewiesen, unter der er von den zu befragenden Nutzern erreichbar ist. Alternativ wäre eine dynamische Einbindung des bestehenden Bogens auf der Website des Kunden denkbar. Nachdem der Bogen veröffentlicht wurde, ist keine Änderung an Itemzahl oder Itemanordnung mehr möglich.

### *Auswertung erhobener Daten aus einem Fragebogen*

Der Nutzer hat die Möglichkeit, Daten, die mit einem bestehenden Bogen erhoben wurden, automatisch auswerten zu lassen. Die Auswertung erfolgt für jedes Item einzeln, da die angefragten Funktionen zum größten Teil voneinander unabhängig sind. Die Visualisierung wurde mithilfe kolorierter Balken vorgenommen, um eine intuitive Wahrnehmung des jeweiligen Werts zu gewährleisten.

Die Likert-Aussagen des Fragebogens sind größtenteils positiv formuliert. Ihre Polung beläuft sich auf 1 (stimme nicht zu) bis 5 (stimme vollkommen zu). Bei entsprechender Umpolung kann ein hoher Summenwert aller Antworten pro Item als positiv visualisiert werden, wohingegen ein niedriger Summenwert eine schlechtere Gebrauchstauglichkeit indiziert. Ein Gesamtwert über alle Items soll zunächst nicht gebildet werden, da es hierbei möglich wäre, dass besonders positiv bewertete Funktionen schlechtere Bewertungen kompensieren. Dies würde entsprechend das Ergebnis der Evaluation verfälschen. Ein weiterer Vorteil einer einzelnen Auswertung der Items ist, dass der Nutzer auf diese Weise einfach und intuitiv erkennen kann, welche Elemente seiner Website Schwachstellen aufweisen.

### *Umfassende aktive und passive Hilfestellung*

Da beim Nutzer keine Kenntnisse im Bereich Fragebogenkonstruktion oder Usability vorausgesetzt werden können, ist vor allem eine umfassende Hilfestellung von Bedeutung. Diese erfolgt sowohl passiv (Bereitstellung eines umfassenden Handbuchs) als auch aktiv (Hinweise zu Bogenerstellung und Bogenlänge während der Bearbeitung).

## 11.4 AKTIVITÄTSDIAGRAMM

Um die einzelnen Funktionen und den Ablauf des Programmes besser visualisieren zu können, sollen im Folgenden die oben genannten Schritte in einem Aktivitätsdiagramm wiedergegeben werden. Bei einem Aktivitätsdiagramm handelt es sich um eine UML-Modellierungsform, mit der funktionale Strukturen eines Programms oder Systems visualisiert werden können (Ablaufdiagramme), wobei der Schwerpunkt der Visualisierung auf den Ablaufmöglichkeiten und -bedingungen unterschiedlicher Aktionen (Tätigkeiten) liegt. Die zeitlich streng sequenzielle Abfolge unterschiedlicher Tätigkeiten stellt hierbei eine Möglichkeit, keine Bedingung dar [Balzert 2009:88ff.].

Das Aktivitätsdiagramm kann sowohl horizontal als auch vertikal dargestellt werden. Die folgende Abbildung zeigt die Abfolge des Programmes aus Gründen des Leseflusses vom Start bis zum Ende in vertikaler Ausrichtung. Hierbei werden die im oberen Abschnitt festgelegten Funktionen berücksichtigt und in ihrer Ablauffolge dargestellt. Rückwärtsgerichtete Pfeile (von unten nach oben) visualisieren hierbei Rückschritte, Wiederholungen oder Sprünge an den Anfang des Programms. Da die Erstellung eines Fragebogens kein streng sequenzieller Vorgang ist, sondern mit hoher Wahrscheinlichkeit Überarbeitungen und Änderungen vor allem bei der Itemzusammenstellung nach sich zieht, verdeutlicht die Darstellung innerhalb des Aktivitätsdiagrammes den Programmablauf am besten.

*Aktivitäten* werden hierbei gemäß der UML-Notation als abgerundete Rechtecke dargestellt. Sie bilden den Aufruf einer neuen Benutzeroberfläche ab, auf der der Nutzer sogenannte *Aktionen* (als beschriftete Pfeile dargestellt) ausführen kann, die zu neuen Aktivitäten führen oder zu Rückschritten im Programm bei der Wiederholung bereits absolvierter Arbeitsschritte. *Entscheidungsknoten* (dichotom/polytom) werden als *Rauten* mit einem eingehenden und mehreren ausgehenden Knoten dargestellt. Hier bieten sich dem Benutzer zwei oder mehr Möglichkeiten für durchzuführende Aktivitäten. Von Entscheidungsrauten zu unterscheiden sind *Vereinigungsknoten* (sog. *merging knots*), die als Rauten mit mehreren eingehenden aber nur einer ausgehenden Kante dargestellt werden. Im folgenden Diagramm werden sie auch als Hilfsknoten verwendet, wenn zwei Aktivitäten, die semantisch voneinander unabhängig sind, dieselbe Konsequenz nach sich ziehen.




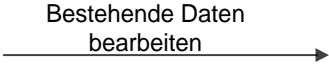
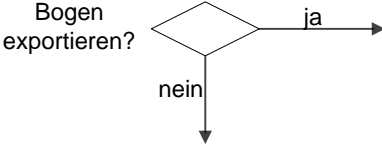
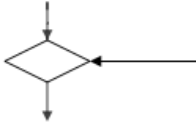


Notation	Abbildung	Bedeutung
<b>Aktivität</b>		Neue Benutzeroberfläche
<b>Aktion</b>		Ausführbare Tätigkeit
<b>Entscheidungsknoten</b>		Mehrere Möglichkeiten für Aktivitäten/Aktionen
<b>Vereinigungsknoten</b>		Zwei oder mehr Aktivitäten/Aktionen ziehen dieselbe Aktion/Tätigkeit nach sich
<b>Startpunkt</b>		Start des Programm
<b>Endpunkt</b>		Ende des Programms

Tabelle 40: Legende des Aktivitätsdiagramms

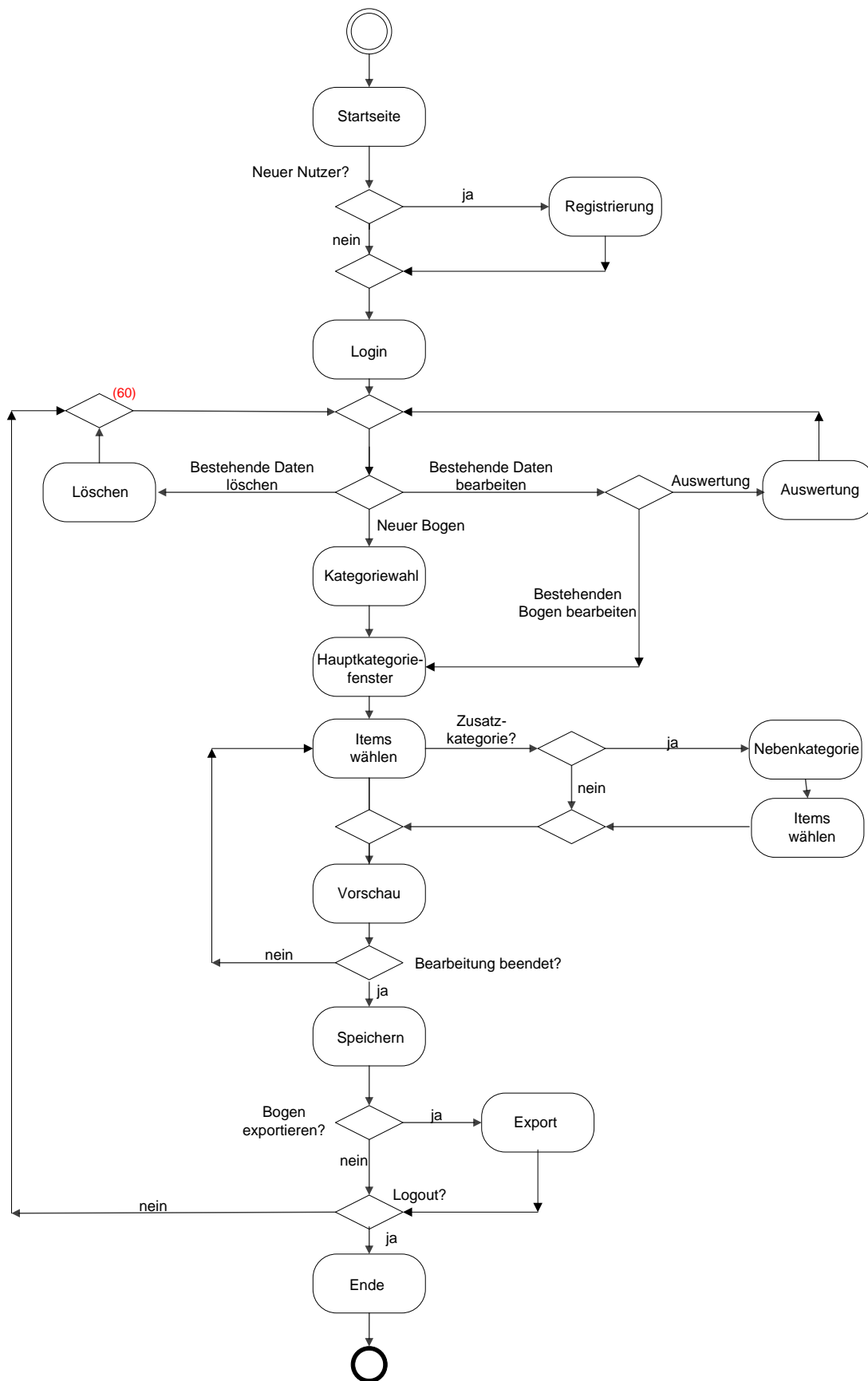


Abbildung 17: Aktivitätsdiagramm des Fragebogengenerators

<sup>60</sup> Der Vereinigungsknoten stellt an dieser Stelle lediglich einen Hilfsknoten zur Vereinigung zweier Kanten mit dem gleichen Ziel dar.

## 11.5 GESTALTGRUNDSÄTZE NACH ISO-9241 FÜR DEN PROTOTYP

Nachdem im vorangegangenen Schritt die Funktionalitäten und der Ablauf des Programms weitestgehend festgelegt wurden, soll im Folgenden die Benutzeroberfläche der Anwendung entworfen werden. Dabei soll vor allem auf Übersichtlichkeit und eine klare Strukturierung der Oberfläche geachtet werden, um auch unerfahrenen Nutzern den Einstieg in das Programm möglichst einfach und selbsterklärend zu gestalten.

Auf die Grundlagen der Gestaltung von Dialogen nach ISO-9241/10 wurde in Abschnitt 7.4 bereits eingegangen. Diese fließen folgendermaßen in die zu erstellende Benutzeroberfläche einfließen:

### *Aufgabenangemessenheit:*

Dem Benutzer wird es ermöglicht, in angemessener Zeit einen Fragebogen gemäß seinen Wünschen zu erstellen und online zu verwenden. Zu diesem Zweck stehen mehrere Itemsätze und -pakete zur Verfügung, die dem Nutzer präsentiert werden, und aus denen er den Bogen schnell, flexibel und möglichst einfach zusammenstellen kann. Zusätzlich wird dem Nutzer zur Erstellung seines Fragebogens auf Wunsch als Orientierung ein bearbeitbarer Standard-Fragebogen mit den wichtigsten Items präsentiert. Die Auswertung sollte möglichst einfach zu erfassen sein und Ergebnisse mithilfe von Grafiken visualisieren.

### *Selbstbeschreibungsfähigkeit*

Dem Benutzer sollte jederzeit klar sein, welche Funktionen das Programm bietet. Da die Anzahl der Funktionen des Programmes überschaubar ist (>10 aktive Funktionen<sup>61</sup>), werden alle Funktionen, die zum Programmzeitpunkt zur Verfügung stehen, im Bildschirm deutlich und dauerhaft ikonisch und verbal visualisiert (beschriftete Buttons mit Abbildungen und Tooltips). Hinweise oder Rückmeldungen des Systems bei fehlerhaften oder fehlenden Eingaben oder erwarteten Eingaben sind möglichst einfach, kurz und verständlich gehalten.

### *Steuerbarkeit*

Der Benutzer hat die Möglichkeit, die Arbeit mit dem Fragebogengenerator jederzeit zu unterbrechen, und zu späterem Zeitpunkt wieder aufzunehmen. Hierzu steht die „Speichern“-Funktion zur Verfügung. Weiterhin ist jederzeit die Möglichkeit zu Abbruch oder Wechsel der aktuellen Aufgabe möglich.

### *Erwartungskonformität*

Hierbei wurde vor allem auf die Konsistenz der Darstellung der Benutzeroberfläche geachtet. Buttons sind einheitlich und „sprechend“ beschriftet, damit der Benutzer sich eine Vorstellung von der Aktion machen kann, die die Interaktion mit Elementen der Benutzeroberfläche auslöst.

---

<sup>61</sup> Registrierung und Login werden einmalig vorgenommen, Die Funktion „Daten auswerten“ stellt einen eigenen Bereich des Programms dar und beinhaltet lediglich die weiteren übergreifenden Funktionen „Hilfe“ und „Abbrechen“ sowie evtl. „Ergebnisse exportieren“

Dem Nutzer werden bei der Kategoriewahl die entsprechenden Items nach Wichtigkeit sortiert präsentiert. Selektion und Deselektion der Items lösen dynamische Veränderungen in der Präsentation der Vorabansicht des Bogens aus. Interaktion über den Browser oder über systeminterne „Zurück“-Buttons führen zur vorhergehenden Seite, nicht an den Anfang des Programms, oder über eine andere (für den Nutzer nicht nachvollziehbare) Seite. Lange Ladezeiten (bei Vorschau oder Export) sollten mit Hilfe einer Grafik visualisiert werden (Fortschrittsbalken, etc.)

#### *Fehlertoleranz*

Bei Eingaben, die unwiderrufliche Konsequenzen haben (Löschen von Datensätzen oder Fragebögen) erfolgt eine Rückmeldung des Systems. Anzudenken ist an dieser Stelle die Einrichtung eines „Papierkorbs“, aus dem Elemente nach dem Löschen erst endgültig entfernt werden können. Weiterhin gibt das System Hinweise, wenn zu viele (kategoriefremde) Items verwendet werden.

#### *Individualisierbarkeit*

Zu überlegen ist, ob es dem Benutzer ermöglicht werden könnte, die Itempräsentation unterschiedlich darstellen zu lassen (verschiedene Benutzeransichten), um dem Nutzer eine Anpassung des Programmes an seine persönlichen Vorlieben anzubieten. Denkbar wäre hier beispielsweise der Wechsel von einer Baum- auf eine Tabellenansicht bei der Präsentation der Items.

#### *Lernförderlichkeit*

Das System stellt dem Benutzer ausreichend aktive (Tooltips, automatisch eingeblendete Hinweise) und passive (umfassende Dokumentation) Hilfestellung zur Verfügung. Wichtig sind vor allem die Erklärung von Haupt- und Nebenkategorien sowie der Hinweis auf die begrenzte Länge des Fragebogens, der mit Hilfe eines Icons in Ampelform vorgenommen wird.

## 11.6 ERSTELLUNG DER BENUTZEROBERFLÄCHE

Nachdem im vorangegangenen Abschnitt die Grundüberlegungen zur Benutzeroberfläche angestellt wurden, konnte der Entwurf für die spätere GUI<sup>62</sup> mit dem Programm AXURE vorgenommen werden. Bei AXURE handelt es sich um einen sogenannten „GUI-Builder“, mit dem innerhalb einer Arbeitsoberfläche (sog. Wireframe) interaktive Elemente (Schaltflächen, Textelemente, etc.) platziert werden und so ein Entwurf für eine Benutzeroberfläche erstellt werden kann (s. Abschnitt 11.7). Das Programm ermöglicht den Export eines interaktiven („klickbaren“) Prototypen, der in einem Standardbrowser bedient wird, und so in der Lage ist, die

---

<sup>62</sup> Graphical User Interface, dt. Benutzeroberfläche

Interaktion mit dem Programm annähernd realitätsgetreu zu simulieren. Dieser Prototyp-Export soll als Grundlage für die Implementierung des späteren Programms dienen.

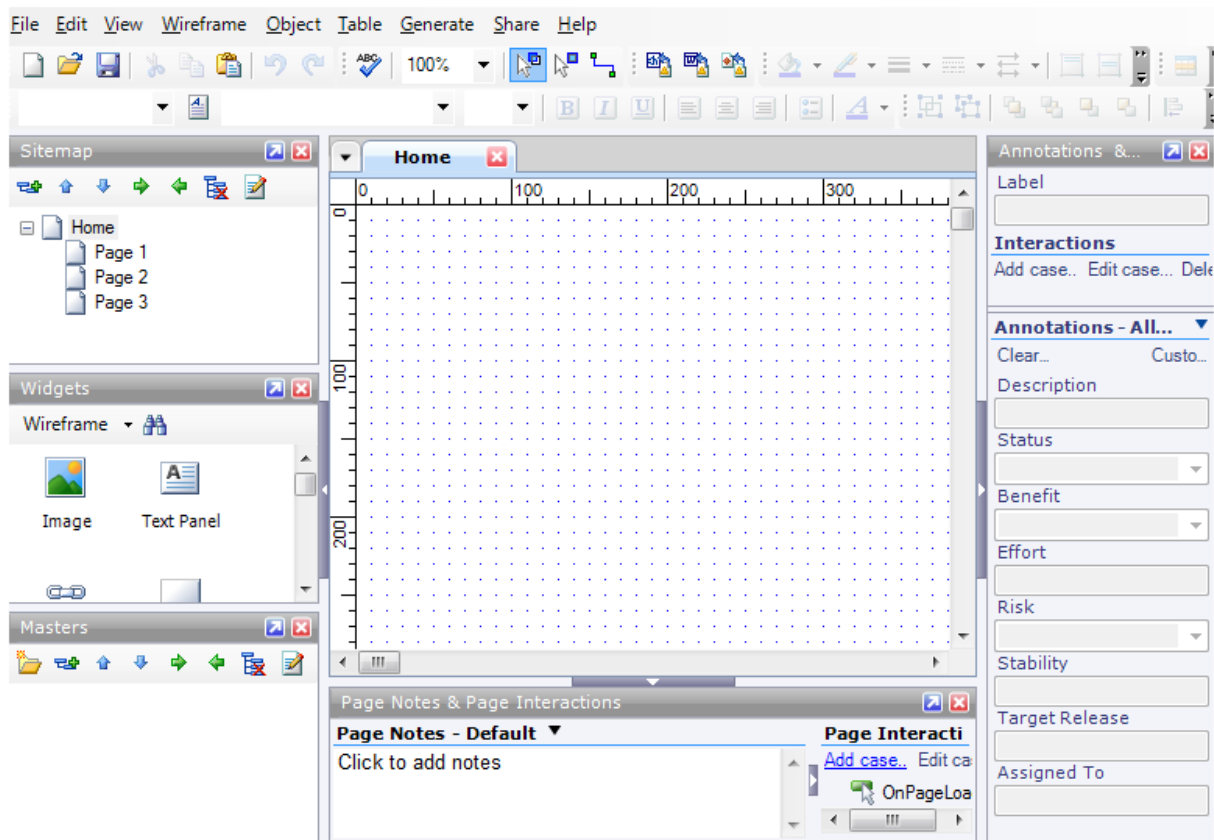


Abbildung 18: Benutzeroberfläche des Programms AXURE

Die Ausfertigung des Prototypen („E-Quest“) geschah in zwei Schritten. Zunächst erfolgte in einer „globalen“ Version die Konzeption der generellen Oberfläche mit der Platzierung von Schaltflächen und Eingabe- sowie Hinweissfeldern. Im zweiten Schritt wurden die Designrichtlinien der Seite [www.usability-toolkit.de](http://www.usability-toolkit.de) in den bestehenden Entwurf eingearbeitet und eine zweite Version erstellt<sup>63</sup>.

Basierend auf diesem Entwurf folgte die Generierung einer „klickbare“ Oberfläche, die als Ansammlung von HTML-Dokumenten, die durch interne Links verknüpft sind und ein begrenztes Ausmaß an Funktionalitäten aufweisen, realisiert wurde. Durch diesen Grad an Interaktivität konnte die spätere Benutzbarkeit des Systems realitätsnah simuliert werden. Der vollständige Prototyp findet sich in Appendix H.

<sup>63</sup> Icons und Logos, die nicht innerhalb von usability-toolkit zur Verfügung standen, wurden von der Seite [www.freeiconsdownload.com](http://www.freeiconsdownload.com) geladen und zum Teil abgeändert. Sie unterliegen der *Creative Commons Attribution 3.0 License* und können zum nichtkommerziellen Gebrauch frei verwendet werden.

## 11.7 SEITENSTRUKTUR DES PROTOTYPEN

Der Prototyp besteht aus insgesamt 53 HTML-Seiten um eine ausreichende Funktionalität zu gewährleisten. Konstruiert wurden exemplarisch die Benutzungsmöglichkeiten „Neuen Bogen erzeugen“ sowie „Bestehenden Bogen bearbeiten“ für die Kategorien E-Commerce, Entertainment und Social Software. Zusätzlich erfolgte die Erstellung von Hinweis- und Hilfeseiten. Für jeden Bereich (Neuer Bogen, E-Commerce, Entertainment, Social Software) wurden eine Übersichts- und eine Hauptseite sowie Seiten für Items von Zusatzkategorien, Vorschau, Auswertung und Übersichtsdaten angelegt.

Seiten mit Hauptfunktion	Hilfe-und Zusatzseiten
1) Home	28) Hilfe
2) Willkommen	29) Registrierung
3) Neuer Bogen	30) Registrierung Bestätigung
4) Default_Main	31) Löschen
5) Daten Default	32) LöschenEC
6) Übersicht_Default	33) LöschenET
7) ZusätzeDefault	34) LöschenSS
8) Vorschau_Default	35) Löschen_Bestätigung
9) Auswertung_Default	36) Export
10) E-Commerce	37) Export_Bestätigung
11) Daten EC	38) Speichern
12) Übersicht_E-Commerce	39) Speichern_Bestätigung
13) E-Commerce_Zusätze	40) Speichern_BestätigungEC
14) Vorschau_E-Commerce	41) Speichern_BestätigungET
15) Auswertung E-Commerce	42) Speichern_BestätigungSS
16) Entertainment	43) Beenden
17) Daten ET	44) BeendenEC
18) Übersicht Entertainment	45) BeendenET
19) Entertainment_Zusätze	46) BeendenSS
20) Entertainment_Vorschau	47) Default
21) Auswertung Entertainment	48) Übersicht-bestätigung_default
22) Social Software	49) Übersicht-bestätigung_EC
23) Daten SS	50) Übersicht-bestätigung_ET
24) Übersicht_Social Software	51) Übersicht-bestätigung_SS
25) Social Software_Zusätze	52) Logout
26) Social Software_Vorschau	53) Logout Bestätigung
27) Auswertung Social Software	

Tabelle 41: HTML-Seiten des erstellten Prototyps

Im folgenden Abschnitt sollen einige Elemente der Benutzeroberfläche kurz beschrieben und erläutert werden, um die Arbeitsweise mit dem Prototypen und dessen Bedienbarkeit zu veranschaulichen.

### 11.7.1 STARTSEITE

Die Startseite (Home.html) soll sich vor allem durch Übersichtlichkeit auszeichnen und beinhaltet eine Registrierungsfunktion und einen Login.

Abbildung 19: E-Quest – Startseite

Sämtliche Buttons (hier: Login und Registrierung) sind, wie in der nachfolgenden Tabelle exemplarisch dargestellt, mit Links zu den entsprechenden Seiten belegt:


Button	Interaktion
1	OnClick:  Open Registrierung in Current Window
2	OnClick:  Open Willkommen in Current Window

Tabelle 42: Verlinkung der HTML-Seite auf den Buttons

### 11.7.2 BEGRÜßUNG

Nach dem Login befindet sich der Nutzer auf der allgemeinen Startseite des Programms (Willkommen.html), auf der er wählen kann, ob er einen bestehenden Bogen bearbeiten oder einen neuen Fragebogen anlegen möchte. Bereits angelegte Bögen des Nutzers können hierbei über ein Dropdown-Menü ausgewählt werden. Die Funktionen „Hilfe“ und „Logout“ stehen auf jeder Seite zur Verfügung. Buttons und Icons sind mit Tooltips versehen, zusätzlich finden sich, falls nötig, kurze Erklärungen auf jeder Seite.

Eingeloggt als:  
**{Vorname, Nachname}**



**E-Quest**  
**Herzlich Willkommen, {Vorname, Nachname}!**  
Bitte wählen Sie unten aus dem Menü aus,  
was Sie heute tun möchten!

Neu hier?  
Besuchen Sie unsere [Hilfe](#)

**Neuen Bogen erstellen:**

Mit dieser Option können Sie einen neuen Fragebogen in einer von Ihnen gewählten Kategorie zusammenstellen und veröffentlichen.

**Bestehenden Bogen aufrufen:**

Bitte wählen Sie hier die Kategorie aus, die am ehesten Ihrer Website entspricht. Wenn Sie z.B. einen Online-Shop besitzen, wählen Sie bitte E-Commerce.

Abbildung 20: E-Quest – Begrüßungsbildschirm nach dem Login



### 11.7.3 NEUEN BOGEN ERSTELLEN

Im Falle der Entscheidung für den Entwurf eines neuen Bogens wird der Nutzer auf eine Seite weitergeleitet, auf der er auswählen kann, für welche Kategorie der neue Bogen entworfen werden soll. Die Kategorienwahl ist wiederum als Dropdown-Menü realisiert. Erklärungen zur Haupt- und Zusatzkategorien finden sich jeweils auf der Seite selber und in Tooltips.

Eingeloggt als:  
{Vorname, Nachname}

**Neuen Bogen erstellen:**

Bitte wählen Sie eine Kategorie aus, die Ihrer Website am ehesten entspricht!

Hauptkategorie: - Kategorie wählen -

Bitte wählen Sie hier die Kategorie aus, die am ehesten Ihrer Website entspricht.  
Wenn Sie z.B. einen Online-Shop besitzen, wählen Sie bitte E-Commerce.

Hinweis: Zur Abfrage zusätzlicher oder von der Hauptkategorie abweichender Funktionalitäten Ihrer Website können Sie im Hauptmenü die Registerkarte "Zusatzkategorien" auswählen!

Los!

Zurück

Hilfe

Log out

Abbildung 21: E-Quest – Erstellung eines neuen Bogens und Kategoriewahl

Nach Auswahl der Hauptkategorie gelangt der Nutzer auf die entsprechende Seite mit den kategoriespezifischen Items. Hier ist die Selektion einer beliebigen Anzahl Items möglich, die dann zu einem Fragebogen zusammengestellt werden. Abhängig von der Zahl ausgewählter Items wird im unteren Bereich der Benutzeroberfläche ein Button grün (Anzahl von Items in Ordnung), gelb (Anzahl der Items kritisch) oder rot (Anzahl der Items zu hoch) gefärbt, da die Zahl der Items 15 – 20 nicht überschreiten sollte.

#### 11.7.4 ITEMS DER HAUPTKATEGORIE

Zur besseren Übersichtlichkeit sind die Aussagen als Baumstruktur angeordnet, so dass der Nutzer einzelne Untergruppen auf- und zuklappen kann. Die Items sind funktional gegliedert, um die Bedienung intuitiver für den Nutzer zu gestalten. Im vorliegenden Beispiel<sup>64</sup> finden sich die Unterkategorien „Seitenaufbau“, „Hilfe“, „Suche“, Funktionen“ und „Multimedia“. Die Items können per Checkbox selektiert/deselektiert werden, ihre Relevanz wird dem Nutzer durch die Anzahl an Sternen visualisiert<sup>65</sup>. Die Möglichkeit, Hinweise oder Erklärungen zu einzelnen Items einzutragen ließ sich mit Axure nicht realisieren, ist aber als Eintrag innerhalb eines Kontextmenüs für jedes Item denkbar. Ein entsprechender Hinweis sollte dann im oberen Bereich der Seite platziert werden.

Abbildung 22: E-Quest – Items der Hauptkategorie

<sup>64</sup> Es handelt sich bei den angezeigten Items nicht um eine bestimmte Kategorie. Da bei der Interaktion mit einem neu zu erstellenden Bogen sechs unterschiedliche Kategorien ausgewählt werden können, wurden aus Umfangsgründen ausgewählte Itempakete zusammengestellt, um die spätere Oberfläche zu simulieren.

<sup>65</sup> Im vorliegenden Prototyp wurden die Checkboxes und Bewertungsvisualisierungen (Sterne) nur auf der Seite Default\_Main.html eingefügt. Da Axure innerhalb einer Baumstruktur nicht die automatische Ein- und Ausblendung von zusätzlichen Elementen (Checkbocnen, Icons) beim Ein- und Ausklappen zulässt, wurde dies hier exemplarisch veranschaulicht und fehlt bei den allen anderen Items.

Die Anzahl gewählter Items wird dem Nutzer an insgesamt drei Stellen angezeigt: Zum einen in der horizontalen oberen Reiterleiste, zum zweiten direkt in der Benutzeroberfläche als Gesamtzahl mit Ampelvisualisierung und zum dritten in der Baumstruktur der Items (siehe auch Abbildung 23, da in der Default\_Main.html standardmäßig keine Items ausgewählt sind). Der Nutzer hat ebenfalls die Möglichkeit, ein vordefiniertes Set von Items automatisch erzeugen zu lassen. Mit der Schaltfläche „Auf Standardauswahl setzen“ wird – nach einem Dialogfeld mit Rückfrage – ein bestimmter Satz von Items zu einem Fragebogen zusammengestellt. Da die möglicherweise bereits getätigten Einstellungen des Benutzers hierbei überschrieben werden, ist die Rückfrage an dieser Stelle essentiell wichtig.

In der oberen horizontalen Reiternavigation ist es dem Benutzer möglich, die Seite mit den Items der Zusatzkategorien auszuwählen, allgemeine Daten und Einstellungen für seinen Bogen zu speichern, sowie sich die Vorschau des erzeugten Bogens anzeigen zu lassen. Weiterhin stehen dem Nutzer auf der Seite die Möglichkeiten zur Vorschau, zum Speichern und zum Beenden des Programms zur Verfügung.

#### 11.7.5 ZUSATZKATEGORIEN

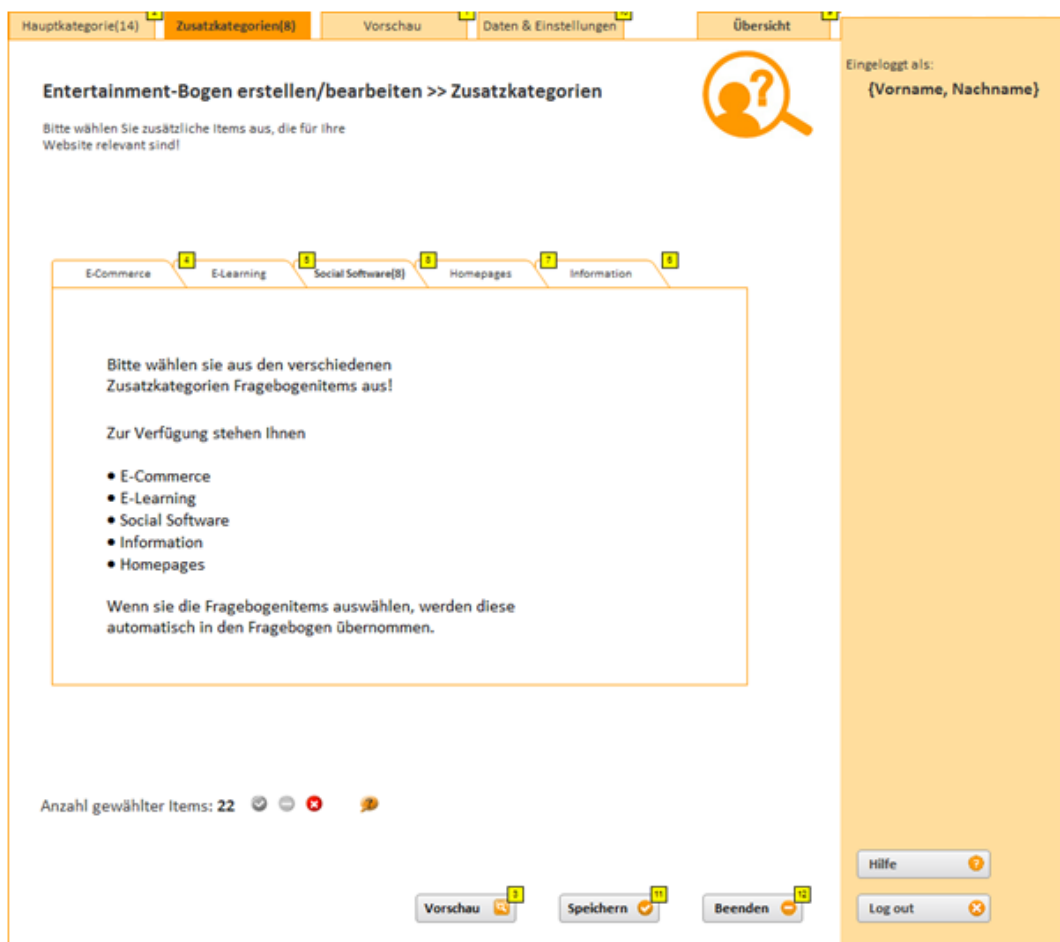


Abbildung 23: E-Quest – Items der Zusatzkategorie

Im Bereich Zusatzkategorien hat der Nutzer die Möglichkeit, sogenannte „kategoriefremde“ Items zu seinem Fragebogen hinzuzufügen, also Items, die keine Funktionen der Hauptkategorie abfragen, sondern einer anderen Kategorie zuzuordnen sind. Nachfolgend beispielhaft die Zusatzitems bei der Wahl der Kategorie Entertainment.

Die verschiedenen Zusatzkategorien mit den jeweiligen Items werden im mittleren Bereich der Oberfläche ebenfalls über eine Reiternavigation angeboten, in der – analog zur oberen Reiternavigation – jederzeit die Anzahl der gewählten Items visualisiert wird. Die Konsistenz der Benutzeroberfläche soll auf diese Art und Weise gewährleistet werden. Wie in Abbildung 24 verdeutlicht, sind in der Kategorie Entertainment insgesamt 22 Items ausgewählt (zu sehen in der Reiternavigation und im Bild). Da dies über die empfohlene Maximalzahl von 15-20 Items pro Bogen hinausgeht, wird im unteren Teil des Bildes ein Hinweis in Form einer roten Ampelfarbe erzeugt.

#### 11.7.6 VORSCHAU

Nachfolgend beispielhaft die Vorschau des bestehenden Bogens der Kategorie „Social Software“<sup>66</sup>:

	Hauptkategorie
1	Die Funktionen der Seite sind auf Anhieb ersichtlich
2	Die Funktionen zur Kommunikation mit anderen Nutzern sind leicht zu bedienen
3	Die Möglichkeiten, mit anderen Nutzern zu kommunizieren, sind ausreichend
4	Der Texteditor ist einfach zu bedienen
5	Ich bin mir sicher, dass mit meinen privaten Daten verantwortlich umgegangen wird
6	Ich würde mein Nutzerprofil gern ausführlicher beschreiben
7	Es ist mir möglich, nur so viele Informationen über mich preiszugeben, wie ich es möchte
8	Ich fühle mich durch eingeschränkte Sichtbarkeit meines Profils gegenüber Dritten sicherer
9	Die Hilfefunktion ist jederzeit auffindbar
10	Die Hilfefunktion ist ausreichend für mich
11	Die Suchfunktion ist leicht zu finden
12	Die Suchfunktion liefert nützliche Ergebnisse
13	Die Benutzeroberfläche ist übersichtlich
14	Das Layout der Seite lässt sich meinen Bedürfnissen entsprechend anpassen
15	Die Möglichkeit zur Bewertung anderer Nutzer ist hilfreich

Anzahl gewählter Items: 15

Zurück Speichern Beenden Hilfe Log out

Abbildung 24: E-Quest – Vorschau des Social Software-Bogens

<sup>66</sup> Die Likert-Skalen wurden aus Platzgründen zunächst weggelassen, sollen jedoch später in der Auswertung erscheinen. Hier ist ebenfalls die Vorschau auf den erstellten Bogen innerhalb eines Popups denkbar.

Im Bereich Vorschau hat der Nutzer eine Ansicht der von ihm zusammengestellten Items. Auch die Möglichkeit zur Änderung der Itemreihenfolge besteht an dieser Stelle.

#### 11.7.7 DATEN UND EINSTELLUNGEN

Abbildung 25: E-Quest – Daten und Einstellungen

Im Bereich Daten und Einstellungen ist es dem Benutzer möglich, einen Titel für den Fragebogen festzulegen, Begrüßungs- sowie Verabschiedungstexte festzulegen, sowie eine Datenschutzerklärung abzufassen. Auch die URL der zu evaluierenden Website kann hier extra eingetragen werden, sofern sie nicht im Begrüßungstext stehen soll. Es besteht die Möglichkeit, die Umfrage nur einem eingeschränkten Teilnehmerkreis zugänglich zu machen, dies geschieht mit Hilfe eines Zugangsschlüssels. Ebenso kann ein Freitextfeld am Schluss für die Speicherung der E-Mail-Adresse, Fragen, Anregungen oder Kommentare mit veröffentlicht werden, wenn gewünscht. Weiterhin besteht die Möglichkeit, alle Fragen als obligatorisch zu kennzeichnen.

## 11.7.8 AUSWERTUNG

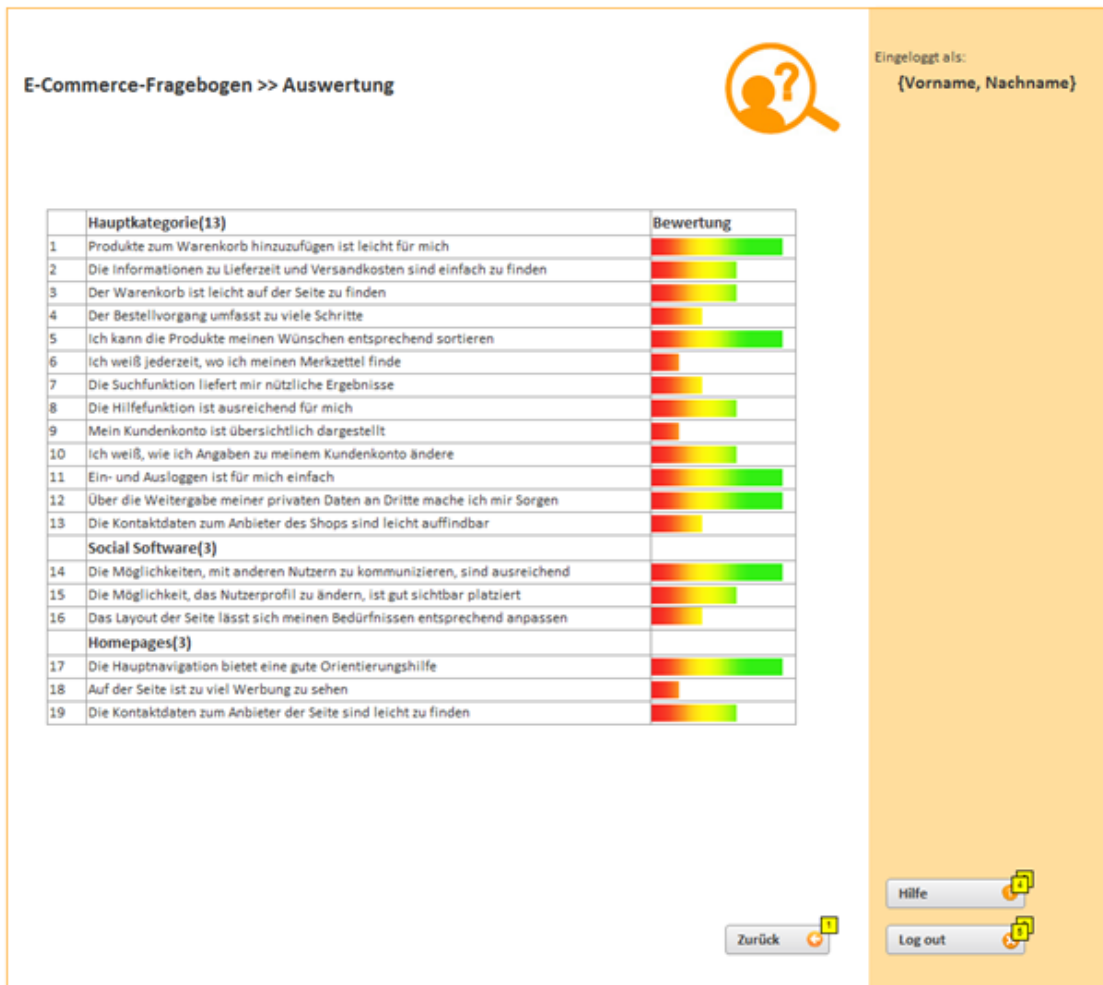


Abbildung 26: E-Quest – Auswertung der Kategorie E-Commerce

Die Auswertung soll mithilfe einer grafischen Visualisierung die Ergebnisse des durchgeführten Fragebogens einfach und intuitiv für den Benutzer darstellen. Hierbei wird ein Balken verwendet, der je nach positiver oder negativer Aussage farblich gestaltet ist. So soll auf Anhieb ersichtlich sein, wo sich Schwachstellen der Seite finden. Da die Items der Likert-Skala eine Polung von 1 (stimme nicht zu) bis 5 (stimme voll zu) aufweisen, und überwiegend positiv formuliert sind, könnte anhand des erzielten Rohwerts (Durchschnitt eines Itemergebnisses über alle Teilnehmer) der Balken erstellt werden. Hierzu müssten dann Items mit negativer Formulierung im Vorfeld innerhalb des Systems entsprechend umgepolt sein.

## 11.8 ZUSAMMENFASSUNG

Unter Einbeziehung des entwickelten Ablaufdiagrammes, der erstellten Anforderungsanalyse sowie der Gestaltrichtlinien der Plattform [www.usability-toolkit.de](http://www.usability-toolkit.de) wurde im vorangegangenen Abschnitt der prototypische Entwurf des späteren Fragebogens durchgeführt. Das Tool soll dazu dienen, dem späteren Nutzer auf eine intuitive Art und Weise die Erstellung des Evaluationsfragebogens für seine Website zu ermöglichen. Von kleineren Einschränkungen, die auf das Programm Axure zurückzuführen sind, abgesehen, konnte eine ausreichend interaktive Oberfläche erstellt werden, die in der Lage ist, die Arbeit mit der Anwendung realitätsnah zu simulieren. Sie bildet ebenfalls die Grundlage für eine, außerhalb des Rahmens dieser Arbeit stattfindende, spätere Implementierung.

## 12 FAZIT UND AUSBLICK

Ziel der vorliegenden Arbeit war es, ein Tool zur Erstellung kontextspezifisch adaptierbarer Fragebögen zur Erhebung der Gebrauchstauglichkeit von Websites zu erstellen. Die Anwendung, mit der dies möglich sein soll, wird auf der Plattform [www.usability-toolkit.de](http://www.usability-toolkit.de) zur Verfügung stehen. Sie soll es Nutzern ermöglichen, ohne Fachkenntnisse oder Vorwissen im Bereich Usability-Evaluation einen Benutzerfragebogen, der auf den Kontext der jeweiligen Website zugeschnitten ist, zu erstellen.

Hierzu wurde zunächst, nach einer allgemeinen Einführung in das Thema Web-Usability und Fragebogenkonstruktion, eine Klassifikation unterschiedlicher Anwendungskontexte des World Wide Web erstellt. Schwierigkeiten, die sich hierbei ergaben, waren vor allem in der dynamischen und multidimensionalen Natur des World Wide Web begründet. Eine unstrittige und strikte Kategorisierung einzelner Websites gestaltet sich entsprechend als schwierig, es war jedoch möglich, Kategorien zu erstellen, in denen die meisten Websites anhand ihrer vorherrschenden Funktion eingeordnet werden können. Diese beschriebene Eigenschaft des World Wide Web allein bringt bereits die Notwendigkeit adaptierbarer Fragebögen mit sich.

Anhand der erstellten Klassifikation erfolgte im Anschluss die Recherche von spezifischen Usability-Kriterien aus der Literatur. Auf die Miteinbeziehung generischer (allgemeiner) Usability-Kriterien, wie sie zur Evaluation von Webauftritten bereits zahlreich vorliegen, wurde in der vorliegenden Arbeit bewusst verzichtet. Der Grund dafür liegt zum einen im begrenzten Umfang des Nutzerfragebogens im Web begründet: Die Evaluation von allgemeinen und kategoriespezifischen Usability-Kriterien würde den Rahmen des Bogen überschreiten. Zum anderen könnte die allgemeine Usability einer Seite beispielsweise im Rahmen einer Expertenevaluation erfasst werden.

Aus den ermittelten Kriterien erfolgten im Anschluss die Erstellung von Fragebogenitems und deren Test durch Experten und Studierende. Diese zweifache Evaluation sollte zum einen die

Qualität (Relevanz, Vollständigkeit) und zum anderen ihre Einsatztauglichkeit (Verständlichkeit, Formulierungen) sicherstellen. Die Items wurden gemäß den Anregungen von beiden Teilnehmergruppen überarbeitet, online mit Endnutzern getestet, und die Ergebnisse ausgewertet. Hierbei zeigte sich, dass die Items sich durchweg als gut geeignet für den Einsatz zur Evaluation der jeweiligen Websites zeigen. Als problematisch erwies sich hierbei jedoch, dass (bedingt durch die Anzahl an Kategorien) nur jeweils eine Website mit den Items einer Kategorie evaluiert werden konnte. Daher waren nicht alle Funktionen oder Elemente, die von den Items evaluiert werden sollten, auch auf der Website vorhanden. Dies äußerte sich in einer stark erhöhten Nonresponse-Quote für die jeweiligen Items. Dies zeigte sich besonders deutlich in den Kategorien E-Commerce (bedingt durch Funktionen, die nur nach Nutzerregistrierung zur Verfügung stehen) und E-Learning.

Die Beurteilung der erstellten Items unter den Gesichtspunkten der allgemeinen Gütekriterien zeigte, dass die Items objektiv und inhaltsvalide sind. Zur Absicherung der erzielten Erkenntnisse wären weitere Tests, vor allem auch in Vergleich mit den Resultaten einer Expertenevaluation oder anderen Fragebögen, sinnvoll und interessant.

Der im letzten Abschnitt erstellte Prototyp „E-Quest“ soll den Aufbau und das Design des späteren Fragebogengenerators veranschaulichen. Hierzu wurde unter Zuhilfenahme einer erstellten Anforderungsanalyse und einem Aktivitätsdiagramm ein Prototyp entworfen, der in der Lage ist, begrenzt interaktiv das spätere Programm aber in annähernd allen Funktionalitäten darzustellen. Somit konnte an dieser Stelle die Grundlage für die spätere Implementierung der Anwendung gebildet werden.

Da mit zunehmender Globalisierung und Technisierung die Verbreitung des World Wide Web immer mehr ansteigt, und sich ein immer größerer Teil des menschlichen Lebens in den Web-Kontext verlagert, wird auch der Faktor Web-Usability immer wichtiger. Die vorliegende Arbeit konnte somit durch die Erstellung kategoriespezifischer Usability-Fragebogenitems einen wichtigen Beitrag zur Ergänzung bereits bestehender Methoden der Usability-Evaluation und der Verbesserung von Webauftritten leisten. Da es sich bei der Web-Usability-Forschung um ein dynamisches Forschungsfeld handelt, besteht die Option, die bestehenden Itemsets zukünftig noch zu ergänzen und zu erweitern.



- BALZERT (2009): UML 2 in 5 Tagen. Der schnelle Einstieg in die Objektorientierung. 2. Auflage. Bochum : W3L GmbH
- BORTZ, DÖRING (2006): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 4., überarbeitete Auflage. Heidelberg : Springer Medizin
- BÜHNER (2006): Einführung in die Test- und Fragebogenkonstruktion. 2., aktualisierte Auflage. München : Pearson Studium
- DOHMANN (2002): Die Praxis des E-Business: Technische, betriebswirtschaftliche und rechtliche Aspekte. Wiesbaden : Vieweg
- GARRETT (2003): The Elements of User Experience. 2. Auflage. Berkeley : New Riders
- GRÄF (2002): Assessing Internet Questionnaires: The Online Pretest Lab. In: Batinić, Reips, Bosnjak, Werner (Hrsg.): Online Social Sciences. S.73-93. Seattle : Hogrefe & Huber
- VON GIZYCKI, BEIER (2002): Usability. Nutzerfreundliche Websites. Berlin/Heidelberg : Springer
- INTERNATIONAL STANDARDIZATION ORGANIZATION (1996): ISO 9241. Ergonomic requirements for office work with display terminals – part 11: Guidance on usability. Geneva: International Organization for Standardization
- JACKOB, SCHOEN, ZERBACK (Hrsg.)(2009): Sozialforschung im Internet. Wiesbaden : VS Verlag für Sozialwissenschaften
- KANNENGIESSER (2005): Objektorientierte Programmierung mit PHP 5. Poing : Franzis
- MAURER, JANDURA (2009) Masse statt Klasse - eine kritische Anmerkung zu Repräsentativität und Validität von Online-Befragungen. In: Jakob, Schoen, Zerback (Hrsg.): Sozialforschung im Internet. S.61-75. Wiesbaden : VS Verlag für Sozialwissenschaften
- MYERS (2004): Psychologie. Heidelberg : Springer Medizin
- NIELSEN (2000): Der Erfolg des Einfachen. München : Markt + Technik
- PUSCHER (2009): Leitfaden Web-Usability. Heidelberg : dpunkt
- SARODNICK, BRAU (2006): Methoden der Usability-Evaluation. Bern : Huber
- SCHWEIBENZ, THISSEN (2003): Qualität im Web. Berlin/Heidelberg : Springer
- STARKE (2008): Effektive Software-Architekturen: ein praktischer Leitfaden. 3. Auflage. München : Carl Hanser
- THIELSCH (2008): Ästhetik von Websites. Münster : Monsenstein und Vannerdat OHG

## Online-Quellen:

AGOF (2009): Berichtsband zur internet facts 2009-II. Arbeitsgemeinschaft Online-Forschung e.V., Darmstadt (verfügbar unter: <http://www.agof.de/index.566.html>, Zugriff 01.04.2010)

BARGAS-AVILA (2001): Evaluationsmethoden der Mensch-Maschine-Interaktion im eCommerce. Lizentiatsarbeit. Institut für Psychologie - Universität Basel (verfügbar unter: <http://www.psychology.unibas.ch/mmi/downloads/Bargas2001.pdf>, Zugriff 01.04.2010)

BEVAN (1998): Usability Issues in Web Site Design. UPA 1998, Washington DC, USA (verfügbar unter: <http://www.usabilitynet.org/papers/usweb98.pdf>, Zugriff: 01.04.2010)

BUNDESVERBAND DIGITALE WIRTSCHAFT (BVDW) (2008a): Usability - ein wichtiger Erfolgsfaktor für E-Commerce-Webseiten (verfügbar unter: <http://www.bvdw.org/medien/leitfaden-usability-fuer-online-shops?media=51>, Zugriff 01.04.2010)

BUNDESVERBAND DIGITALE WIRTSCHAFT (BVDW) (2008b): Leitfaden Web 2.0 & E-Commerce (verfügbar unter: <http://www.bvdw.org/mybvdw/media/download/bvdw-ecommerce-web20-20070727.pdf?file=30>, Zugriff 01.04.2010)

CHIN, DIEHL, NORMAN (1988): Development of an instrument measuring user satisfaction of the human-computer interface. Proceedings of the SIGCHI conference on Human factors in computing systems. Washington DC, United States. S.213-218 (verfügbar unter: <http://portal.acm.org/citation.cfm?id=57203>, Zugriff: 01.04.2010)

COLLINS (2006): Web Site Appraisal. Google vs. Yahoo. School of Information Resources and Library Sciences (verfügbar unter: [http://www.sir.arizona.edu/iso/bibliotech/2007jan\\_vol4\\_no2/Web\\_Site\\_Appraisal.pdf](http://www.sir.arizona.edu/iso/bibliotech/2007jan_vol4_no2/Web_Site_Appraisal.pdf), Zugriff: 01.04.2010)

deMARSICO, LEVIALDI (2004): Evaluating web sites: exploiting user's expectations. In: Journal of Human-Computer Studies. Ausgabe 60, Nr. 3. S. 381-416 (verfügbar unter: <http://stories.iat.sfu.ca/protected/docs/papers/WebsiteEval.pdf>, Zugriff 01.04.2010)

DESURVIRE, JEGERS, WIBERG (2007): Evaluating Fun and Entertainment: Developing A Conceptual Framework Design of Evaluation Methods. Interact 2007, First Workshop on Design Principles (verfügbar unter: <http://www.fun-of-use.org/interact2007/papers/EvaluatingFunAndEntertainment.pdf>, Zugriff: 01.04.2010)

DESURVIVRE, WIBERG (2009): Game Usability Heuristics (PLAY) For Evaluating and Designing Better Games: The Next Iteration. (verfügbar unter: <http://www.behavioristics.com/downloads/DesurvireFinalHCI09PLAY.pdf>, Zugriff: 01.04.2010)

GEDIGA, HAMBORG, DÜNTSCH (1999). The IsoMetrics Usability Inventory: An operationalisation of ISO 9241-10, In: Behaviour and Information Technology. Ausgabe 18, S. 151 – 164

HAMMWÖHNER (2007): Qualitätsaspekte der Wikipedia. In: kommunikation@gesellschaft. Sonderausgabe: *Wikis - Diskurse, Theorien und Anwendungen* (verfügbar unter: [http://www.soz.uni-frankfurt.de/K.G/B3\\_2007\\_Hammwoehner.pdf](http://www.soz.uni-frankfurt.de/K.G/B3_2007_Hammwoehner.pdf) , Zugriff: 01.04.2010)

HARMS, SCHWEIBENZ, STROBEL (2002): Usability Evaluation von Web-Angeboten mit dem Web Usability Index. In: Proceedings der 24. DGI-Online-Tagung 2002 - Content in Context. Frankfurt/Main (verfügbar unter: [http://usability.is.uni-sb.de/beitrag/web\\_usability\\_index.pdf](http://usability.is.uni-sb.de/beitrag/web_usability_index.pdf), Zugriff: 01.04.2010)

HARPER, NORMAN (1993): Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5. Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference, S. 224-228. (verfügbar unter: <http://www.lap.umd.edu/quis/publications/harper1993.pdf>, Zugriff 01.04.2010)

HASSENZAHL, BURMESTER, KOLLER (2003): AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Mensch und Computer 2003: Interaktion in Bewegung. S. 187-196. Stuttgart : B.G. Teubner

HASSENZAHL, BURMESTER, KOLLER (2008): Der User Experience (UX) auf der Spur: zum Einsatz von [www.attrakdiff.de](http://www.attrakdiff.de). In: Brau, Diefenbach, Hassenzahl, Koller, Peissner,

Röse,(Hrsg.): Usability Professionals 2008: Berichtsband des sechsten Workshops des German Chapters der Usability Professionals Association e.V. S. 78-82. Stuttgart : Fraunhofer IRB

HEIDENREICH, RAFF (2006): Usability ins E-Learning! (verfügbar unter: [http://tu-dresden.de/die\\_tu\\_dresden/zentrale\\_einrichtungen/mdc/aufsaeetze\\_ordner/Usability%20ins%20E-Learning.pdf](http://tu-dresden.de/die_tu_dresden/zentrale_einrichtungen/mdc/aufsaeetze_ordner/Usability%20ins%20E-Learning.pdf), Zugriff 01.04.2010)

HELIC , MAURER, SCERBAKOV (2004): Discussion Forums as Learning Resources in Web-Based Education (verfügbar unter: <http://coronet.iicm.tugraz.at/denis/pubs/ijca2004.pdf>, Zugriff: 01.04.2010)

HOFFMANN, NOVAK, CHATTERJEE (1995): Commercial Scenarios for the Web: Opportunities and Challenges. In: Journal of Computer Mediated Communication. Ausgabe 1, Nr. 3 (verfügbar unter: <http://www.ascusc.org/jcmc/vol1/issue3/hoffman.html>, Zugriff 01.04.2010)

HONG, KIM (2004): Architectural criteria for website evaluation - conceptual framework and empirical validation. In: Behaviour & Information Technology. Ausgabe 23, Nr 5, S. 337-357

KACZMIREK (2008): Human-Survey Interaction. Usability and Nonresponse in Online Surveys. (verfügbar unter: <http://madoc.bib.uni-mannheim.de/madoc/volltexte/2008/2150/pdf/kaczmirek2008.pdf>, Zugriff: 22.03.2010)

KATRE (2007): Model-Based Usability Heuristics for constructivist e-Learning. I-Manager's International Journal of Education Technology. Ausgabe. 3, Nr. 4, S. 58-65 (verfügbar unter: <http://www.hceye.org/downloads/Katre-constructivism-Jan-2007-Imanager.pdf>, Zugriff: 01.04.2010)

KIRAKOWSKI (1994): The Use of Questionnaire Methods for Usability Assessment. Background Notes on the SUMI Questionnaire (verfügbar unter: <http://sumi.ucc.ie/sumipapp.html>, Zugriff 01.04.2010)

KIRSTE , DIENER, ENCARNACAO, HEIN, STORKUND, BIMBER-FRAUNHOFER (2002): Die Zukunft interaktiver Systeme: Herausforderungen an die Usability-Forschung. it+ti - Informationstechnik und Technische Informatik Ausgabe 44(1) S. 40-48.

KOCH (2008): Enterprise 2.0. Social Software in Unternehmen. White Paper, Universität der Bundeswehr München (verfügbar unter: [http://www.kooperationssysteme.de/docs/pubs/Koch2008-bericht\\_enterprise20.pdf](http://www.kooperationssysteme.de/docs/pubs/Koch2008-bericht_enterprise20.pdf), Zugriff: 01.04.2010)

KUKULSKA-HULME (2004): Usability: A Common Theme for Developments in e-Learning at the UK Open University (verfügbar unter: [www.igi-pub.com/files/prefaces/jdet%20preface%20\(3\).pdf](http://www.igi-pub.com/files/prefaces/jdet%20preface%20(3).pdf), Zugriff: 23.Dezember.2009)

LATTEN (2001): Utility und Usability von E-Learning. Forschungsgruppe Instruktion und interaktive Medien (IIM). Universität Gießen (verfügbar unter: [http://www.uni-saarland.de/fileadmin/user\\_upload/extranet\\_uds/campus/struktur/sonstig/competencecenter\\_virtuelesaar/Veranstaltungen/2001/ppt\\_latten.pdf](http://www.uni-saarland.de/fileadmin/user_upload/extranet_uds/campus/struktur/sonstig/competencecenter_virtuelesaar/Veranstaltungen/2001/ppt_latten.pdf), Zugriff: 01.04.2010)

LAUGWITZ, SCHREPP, HELD (2008): Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten In: Heinecke, Paul (Hrsg.): Mensch & Computer 2006: Mensch und Computer im StrukturWandel. S.125-137. München : Oldenbourg

LIU, MOTURU (2009): Evaluating the Trustworthiness of Wikipedia Articles through Quality and Credibility. International Symposium on Wikis and Open Collaboration 2009, Orlando, Florida (verfügbar unter: <http://www.wikisym.org/ws2009/procfiles/p127-moturu.pdf>, Zugriff: 01.04.2010)

LORENZO-SEVA, FERRANDO (2005): IRT-Related Factor Analytic Procedures for Testing the Equivalence of Paper-and-Pencil and Internet-Administered Questionnaires. In: Psychological Methods, Ausgabe 10, Nr 2. S. 193-205.

LUDIN (2008): Usability Issues from the NASA and Wikipedia Apollo Sites. (verfügbar unter: [http://www.zarlaludin.com/uploads/1/6/9/7/1697495/issues\\_based\\_metrics.pdf](http://www.zarlaludin.com/uploads/1/6/9/7/1697495/issues_based_metrics.pdf), Zugriff: 01.04.2010)

MANDRYK, INKPEN, CALVERT (2006): Using psychophysiological techniques to measure user experience with entertainment technologies. Preprint. Zu erscheinen in: Behaviour & Information Technology (verfügbar unter: [http://www.reganmandryk.com/pubs/mandryk\\_bit\\_preprint.pdf](http://www.reganmandryk.com/pubs/mandryk_bit_preprint.pdf), Zugriff: 01.04.2010)

MARIAGE, VANDERDONCKT (2000): A Comparative Usability Study of Electronic Newspapers. ISYS (Information Systems) Research Unit, IAG Louvain School of Management (verfügbar unter: <http://www.isys.ucl.ac.be/bchi/publications/2000/Mariage-TFWWG2000.pdf>, Zugriff: 01.04.2010)

MELIS, WEBER (2003): Lessons for (Pedagogic) Usability of eLearning Systems. In A. Rossett (Hrsg.), Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2003, Phoenix, Arizona. S. 281-284 (verfügbar unter: <http://www-ags.dfki.uni-sb.de/~melis/Pub/elearn03Usability.pdf>)

MÜLLER-PROTHMANN, SIEDENTOPF (2003): Designing Online Knowledge Communities: Developing a Usability Evaluation Criteria Catalogue. 3<sup>rd</sup> European Knowledge Management Summer School. San Sebastian, Spain (verfügbar unter: [http://www.providersedge.com/docs/km\\_articles/Designing\\_Online\\_K\\_Communities\\_-\\_Developing\\_Usability\\_Eval\\_Criteria\\_Catalogue.pdf](http://www.providersedge.com/docs/km_articles/Designing_Online_K_Communities_-_Developing_Usability_Eval_Criteria_Catalogue.pdf), Zugriff: 01.04.2010)

NIELSEN (2001): 113 Design Guidelines for Homepage Usability (verfügbar unter: <http://www.useit.com/homepageusability/guidelines.html>, Zugriff 01.04.2010)

NIELSEN (2003): The Ten Most Violated Homepage Design Guidelines (verfügbar unter: <http://www.useit.com/alertbox/20031110.html>, Zugriff 01.04.2010)

PINHANEZ, KARAT C., KARAT J., VERGO, ARORA, RIECKEN, COFINO (2001): Can Web Entertainment Be Passive? Tenth International World Wide Web Conference (IWWW). Hongkong (verfügbar unter: <http://www10.org/program/society/pinhanez/iwww01.htm>, Zugriff 01.04.2010).

POLL (2007): Evaluating the library website: Statistics and quality measures. World Library and Information Congress: 73<sup>rd</sup> IFLA General Conference and Council. Durban, South Africa (verfügbar unter: <http://archive.ifla.org/IV/ifla73/papers/074-Poll-en.pdf>, Zugriff 01.04.2010)

PREECE (2004): Online Communities: Researching sociability and usability in hard to reach populations. OZCHI 2004. Wollongong University (verfügbar unter: <http://www.ozchi.org/proceedings/2004/pdfs/keynote-preece.pdf>, Zugriff: 01.04.2010)

- PRÜMPER, ANFT (1993): ISONORM 9241/10 (verfügbar unter: <http://www.medpaed.de/dmddocuments/ss07/mumilu/isonorm-fragebogen.pdf>, Zugriff: 22.03.2010)
- RAGANATHAN GANAPATHY (2002): Key dimensions of business-to-consumer web sites. In: Information and Management. Ausgabe 39, Nr. 6. S. 457-465.
- RENTRÓIA-BONITO, GUERREIRO, MARTINS, FERNANDES, JORGE (2006): Evaluating Learning Support Systems Usability. An Empirical Approach. S. 183 - 188 3rd E-Learning Conference, Coimbra, Portugal (verfügbar unter: <http://elconf06.dei.uc.pt/pdfs/paper29.pdf>, Zugriff 01.04.2010)
- SARTZETAKI, PSAROMILIGKOS, RETALIS, AVGERIOU (2003): An approach for usability evaluation of e-commerce sites based on design patterns and heuristics criteria. HCI 2003. Bath, England (verfügbar unter: [http://users.softlab.ece.ntua.gr/~retal/papers/conferences/HCI-Crete/Maria\\_EVAL/HCI2003\\_DEPTH\\_camera.pdf](http://users.softlab.ece.ntua.gr/~retal/papers/conferences/HCI-Crete/Maria_EVAL/HCI2003_DEPTH_camera.pdf), Zugriff 01.04.2010)
- SAUER, SCHMEIßER (2005): Usability-Forschung im E-Commerce. In Planung und Analyse, Ausgabe 1/2005 (verfügbar unter: [http://www.phaydon.de/fileadmin/Bilder/Fachartikel\\_PDF-Versionen/Phaydon\\_Usability.pdf](http://www.phaydon.de/fileadmin/Bilder/Fachartikel_PDF-Versionen/Phaydon_Usability.pdf), Zugriff 01.04.2010)
- SCHNELLER (2009): Zentrale Trends der Internetnutzung in den Bereichen Information, Kommunikation und E-Commerce. Allensbacher Computer- und Technik-Analyse (ACTA) 2009. Institut für Demoskopie Allensbach (verfügbar unter: <http://www.acta-online.de/>, Zugriff 02.04.2010)
- STEWART (2008): Usability or User Experience - What's the difference? (verfügbar unter: <http://www.system-concepts.com/articles/usability-articles/2008/usability-or-user-experience-whats-the-difference.html?>, Zugriff: 01.04.2010)
- TNS Infratest & Initiative D21 (2009): (N)onliner Atlas 2009: Eine Topografie des digitalen Grabens durch Deutschland. (verfügbar unter: <http://www.initiaved21.de/category/nonliner-atlas>, Zugriff 01.04.2010)

WIBERG, JEGERS (2003): Satisfaction and Learnability in Edutainment: A usability study of the knowledge game 'Laser Challenge' at the Nobel e-museum (verfügbar unter: <http://www8.informatik.umu.se/~colsson/cwkjhci03.pdf>, Zugriff: 01.04.2010)

WIBERG (2005): Fun in the Home: Guidelines for Evaluating Interactive Entertainment on the Web (verfügbar unter: <http://www.informatik.umu.se/~colsson/Articles/finalhciint2005CW.pdf>, Zugriff: 01.04.2010)

ZAHARIAS (2004): Developing a Usability Evaluation Method for E-learning Applications: From Functional Usability to Motivation to Learn. PhD Thesis, Athens University of Economics and Business (verfügbar unter: <http://www.eltrun.gr/index.php/phd-studies/completed-phds/340>, Zugriff: 01.04.2010)

ZVIRAN, GLEZER, ITAY (2006): User satisfaction from commercial web sites: The effect of design and use. In: Information and Management. Ausgabe 43, Nr.2. S. 157-178 (verfügbar unter: <http://www.niis.tau.ac.il/admin/media/document/Information%20and%20Management.pdf>, Zugriff 01.04.2010)



## APPENDIX A – GEGENÜBERSTELLUNG DER KATEGORISIERUNGEN VON THIELSCH/HONG & KIM

Thielsch 2008	Hong & Kim 2004
E-Commerce	General Shopping Mall, Specialty Shopping Mall, Auction
Entertainment	Games, Web Casting
E-Learning	Online Education
E-Recruiting/E-Assessment	k.E. <sup>67</sup>
Information	Personal Homepage, Organization Homepage
Kommunikation/Online-Netzwerke	Community
Portalfunktion	Portals
Selbstdarstellung	Personal Homepage, Organization Homepage
Social Software	Community
Suchmaschinen	k.E.
k.E.	Health and Medical
k.E.	Economy and Industry
k.E.	Web Hosting
k.E.	Woman and Children
k.E.	Reservation Service
k.E.	Computer and Internet

<sup>67</sup> keine Entsprechung

## APPENDIX B - ZUORDNUNG DER ERMITTELTEN USABILITY-KRITERIEN ZU DEN EINZELNEN GESTALTGRUNDSÄTZEN

### E-COMMERCE

Grundsätze der Dialoggestaltung	
Aufgabenangemessenheit	<ul style="list-style-type: none"> <li>• Selektive Ansicht durch Facettenklassifikation (Sortierung nach Preis, Datum, etc.)</li> <li>• Informationen über Preis, Lieferzeit, Versandkosten etc.</li> <li>• Kategoriale Auflistung der Produkte</li> <li>• Übersichtlichkeit der linken Navigationsleiste <ul style="list-style-type: none"> <li>○ Begrenzte Anzahl an Produktkategorien</li> <li>○ Begrenzte Anzahl an Ebenen bei kaskadierenden Menüs</li> </ul> </li> <li>• Erreichbarkeit des Warenkorbs von jeder Seite aus</li> <li>• Prominente Platzierung der Suchfunktion</li> <li>• Sichtbare Kontaktdaten zum Anbieter des Shops</li> <li>• Kaufprozess realitätsgetreu abgebildet</li> </ul>
Selbstbeschreibungsfähigkeit	<ul style="list-style-type: none"> <li>• Sichtbare Produktkategorien in der linken und/oder oberen Navigation</li> <li>• Sichtbarkeit der Werkzeuge (Warenkorb, Merkliste)</li> <li>• Prominente Platzierung von Hilfe/Supportfunktion und FAQ</li> <li>• Visualisierung von Lade- und Wartezeiten durch Fortschrittsbalken</li> <li>• Systemrückmeldungen über Erfolg oder Misserfolg bei Eingaben in Formularfelder (z.B. Adressänderung)</li> </ul>
Steuerbarkeit	<ul style="list-style-type: none"> <li>• Items im Warenkorb können zu jedem Zeitpunkt (auch während des Bestellvorgangs) aus dem Warenkorb gelöscht werden</li> <li>• Prominente Platzierung von Hinweisen zu Datensicherheit und Datenschutz</li> <li>• Möglichkeit zur Ein- und Ausblendung bestimmter Bereiche</li> <li>• Wiederherstellung von Eingaben in Formularfeldern bei Betätigung des „Zurück“-Buttons</li> <li>• Wiederherstellung von Kategorieselection oder Facettenklassifikation bei Betätigung des „Zurück“-Buttons</li> <li>• Wiederherstellung von Eingaben in Formularfeldern bei Betätigung des „Zurück“-Buttons</li> <li>• Kategorienübergreifende Vergleichbarkeit</li> </ul>

Erwartungskonformität	<ul style="list-style-type: none"> <li>• Sortierung der Produkte nach Kategorien</li> <li>• Einheitlichkeit Produktdarstellung</li> <li>• Einheitlichkeit von Workflows (Bestell-/Bezahlungsverfahren, Änderungen von Kundendaten)</li> </ul>
Fehlertoleranz	<ul style="list-style-type: none"> <li>• Rechtschreibfehler in der Suche liefern Ergebnisse</li> <li>• Filtereinstellungen liefern auch ähnliche Produkte</li> <li>• Verzicht auf oder deutliche Kennzeichnung des Reset-Buttons (Löschung der Eingaben in Formularfeldern)</li> </ul>
Individualisierbarkeit	<ul style="list-style-type: none"> <li>• denkbar<sup>68</sup>: Passende Produkte zu Produkten im Warenkorb werden angeboten</li> <li>• denkbar: Persönliche Begrüßung</li> <li>• denkbar: Persönliche Empfehlungen</li> </ul>
Lernförderlichkeit	<ul style="list-style-type: none"> <li>• Aktive Hilfefunktionen des Systems bei komplexen Funktionen</li> <li>• (interaktive) Tour durch das System</li> </ul>
Ergänzender Kriterienkatalog	<ul style="list-style-type: none"> <li>• Interessant/Originell gestaltete Website</li> <li>• Interessant gestalteter Ablauf</li> <li>• Platzierung interessanter Produkte (Besteller) auf der Startseite</li> <li>• Schaffen einer Vertrauensbasis durch <ul style="list-style-type: none"> <li>○ Kundenbewertungen des Shops</li> <li>○ Kundenbewertungen/Testberichte der Produkte</li> <li>○ Verkäuferbewertungen (bei Auktionen oder Fremdverkäufern)</li> </ul> </li> <li>• Gütesiegel unabhängiger Prüfer für den Online-Shop</li> <li>• Autovervollständigung während der Eingabe im Suchfeld</li> <li>• Vergrößer- und u.U. drehbare Abbildungen der Produkte</li> <li>• Vermeidung von Einbindung (system-) anspruchsvoller Applets/Plugins</li> <li>• Tagging von Produkten</li> </ul>

<sup>68</sup> eigene Anregung des Autors

Grundsätze der Dialoggestaltung	
Aufgabenangemessenheit	<ul style="list-style-type: none"> <li>• Großzügige Anordnung der Elemente auf dem Bildschirm (Außenabstand)</li> <li>• Begrenzte Anzahl angezeigter Elemente auf dem Bildschirm zur Minimierung des kognitiven Aufwands</li> <li>• Bereitstellung unterschiedlicher Hilfestellungen (FAQ, Tour, Online-Foren)</li> <li>• Wechsel zwischen einzelnen Modulen problemlos möglich</li> </ul>
Selbstbeschreibungsfähigkeit	<ul style="list-style-type: none"> <li>• Hilfefunktionen deutlich sichtbar platziert und von jeder Seite aus abrufbar</li> <li>• Farbliche Gestaltung von Modulen nach Schwierigkeitsgrad</li> <li>• Icons sollten keine komplexen Funktionen beinhalten</li> <li>• Sitemap oder Beschreibung der Struktur des Programms</li> </ul>
Steuerbarkeit	<ul style="list-style-type: none"> <li>• Wiederherstellung von Eingaben in Formularfeldern bei Betätigung des „Zurück“-Buttons</li> <li>• Navigieren zwischen den einzelnen Lernmodulen ist möglich</li> </ul>
Erwartungskonformität	<ul style="list-style-type: none"> <li>• Konsistenz der grafischen Benutzeroberfläche auch bei wechselnden Inhalten</li> </ul>
Fehlertoleranz	<ul style="list-style-type: none"> <li>• Fehlerhafte Eingaben (bspw. bei Übungsaufgaben können korrigiert werden)</li> <li>• Wiederherstellung von bereits absolvierten Inhalten/getätigten Eingaben bei Abbruch</li> </ul>
Individualisierbarkeit	<ul style="list-style-type: none"> <li>• Speicherung bereits vermittelter Lerninhalte</li> <li>• Personalisierte Trainingsmodule mit bereits erlernten Inhalten</li> <li>• Möglichkeit, automatische Hilfefunktionen (Systemhinweise) aus- und einzustellen</li> <li>• Getätigte Einstellungen bleiben auch nach Systemabbruch erhalten</li> </ul>
Lernförderlichkeit	<ul style="list-style-type: none"> <li>• Verschiedene Schwierigkeitsstufen</li> <li>• Lösungen zu Aufgaben werden zunächst durch Hinweise ersetzt und erst im zweiten Durchlauf geliefert</li> </ul>
Ergänzender Kriterienkatalog	
	<ul style="list-style-type: none"> <li>• Unterscheidung von aktiver (Systemhinweise) und passiver Hilfefunktionen (auf Anfrage)</li> </ul>

- Druckfunktionen für lange Texte

## PORTALE/SOCIAL SOFTWARE

Grundsätze der Dialoggestaltung	
Aufgabenangemessenheit	<ul style="list-style-type: none"> <li>• Bereitstellung und sichtbare Platzierung von Funktionen zur Kommunikation <ul style="list-style-type: none"> <li>○ Livechat</li> <li>○ Instant Messenger</li> <li>○ Gästebuch</li> </ul> </li> <li>• Erreichbarkeit der Plattform rund um die Uhr</li> <li>• Änderung des Nutzerprofiles jederzeit möglich</li> </ul>
Selbstbeschreibungsfähigkeit	<ul style="list-style-type: none"> <li>• Deutliche Visualisierung von Ausrichtung und Funktionen des Portals bei Betreten der Seite</li> <li>• Eindeutige Beschriftung von Formularfeldern bei der Profilerstellung</li> <li>• Hinweise auf sog. „Netiquette“</li> <li>• „Tour“ durch das System für neue Nutzer, prominent platzierte und von jeder Seite aus abrufbare Hilfefunktion</li> </ul>
Steuerbarkeit	<ul style="list-style-type: none"> <li>• Gästebucheinträge und Messages können gelöscht werden</li> <li>• Löschen von Einträgen kann rückgängig gemacht werden</li> <li>• Änderung des Nutzerprofiles jederzeit möglich</li> <li>• Einschränkung der Sichtbarkeit des Profils für unterschiedliche Nutzer</li> </ul>
Erwartungskonformität	<ul style="list-style-type: none"> <li>• Anpassung von Editoren an bestehende mentale Modelle</li> <li>• Vorgefertigte Emoticons zur Auswahl</li> <li>• Zielgruppenspezifische Anpassung des Layouts</li> </ul>
Fehlertoleranz	<ul style="list-style-type: none"> <li>• Eindeutige Rückmeldung bei nichterlaubten Bildformaten im Upload (Hinweis auf erlaubte Formate)</li> </ul>
Individualisierbarkeit	<ul style="list-style-type: none"> <li>• Unterschiedliche Detaillierungsgrade des Nutzerprofiles möglich</li> <li>• Anpassbares Layout</li> </ul>
Lernförderlichkeit	<ul style="list-style-type: none"> <li>• Guided Tour durch das System oder interaktive Systemhilfe</li> </ul>
Ergänzender Kriterienkatalog	
	<ul style="list-style-type: none"> <li>• Nutzerbenachrichtigung bei Ausfall der Seite mit voraussichtlicher Downtime</li> <li>• Mitglieder-Ranking/Mitglieder-Bewertungsfunktion</li> <li>• Hinweise auf Datenschutz</li> </ul>

Grundsätze der Dialoggestaltung	
Aufgabenangemessenheit	<ul style="list-style-type: none"> <li>• Aktualität der Informationen</li> <li>• Verständlichkeit der Informationen</li> <li>• Freitext- und Schlagwortsuche</li> <li>• Suchergebnisse nach Aktualität geordnet</li> <li>• Glossar für Sachbegriffe</li> <li>• Linksbündige Ausrichtung des Textes</li> <li>• Vermeidung von serifenhaltigen Schriften zur Verbesserung der Lesbarkeit</li> <li>• Begrenzung der Textlänge (&lt;1,5 Seiten), wichtige Angaben sollten ohne Scrollen lesbar sein</li> <li>• Vermeidung doppelten (vertikal + horizontal) Scrollens</li> <li>• Verwendung kurzer Sätze und zielgruppenspezifischer Wortwahl</li> </ul>
Selbstbeschreibungsfähigkeit	<ul style="list-style-type: none"> <li>• Kurzbeschreibung der Seitengliederung / Kategorisierung auf der Startseite</li> <li>• Inhaltsangaben oder Abstracts am Anfang der textuellen Information</li> </ul>
Kontrollierbarkeit	<ul style="list-style-type: none"> <li>• denkbar: Aus- und einschalten von visuellen oder auditiven Medien</li> </ul>
Erwartungskonformität	<ul style="list-style-type: none"> <li>• Aktuelle Informationen stehen zuoberst</li> <li>• Gliederung von Artikeln: allgemein → speziell</li> <li>• Existenz einer Sitemap zur Orientierung des Nutzers</li> <li>• Alternativtexte für Bilder und Plugins</li> </ul>
Fehlertoleranz	<ul style="list-style-type: none"> <li>• Rechtschreibfehler in der Suche liefern (gültige) Ergebnisse</li> </ul>
Individualisierbarkeit	<ul style="list-style-type: none"> <li>• denkbar: Nutzerprofil mit Funktion zur Archivierung relevanter Information (kürzlich angesehene Artikel)</li> </ul>
Lernförderlichkeit	<ul style="list-style-type: none"> <li>• Gliederung langer Texte in gedankliche Abschnitte (Paragrafen, eine Idee/Absatz)</li> <li>• Zwischenüberschriften zur Übersichtlichkeit</li> </ul>
Ergänzender Kriterienkatalog	
	<ul style="list-style-type: none"> <li>• Datierung der Information</li> <li>• Autorangabe mit Kontaktadresse bei Artikeln</li> <li>• Angebot von Newsletter/RSS-Feeds</li> <li>• Druckfunktionen für lange Texte</li> </ul>

Grundsätze der Dialoggestaltung	
Aufgabenangemessenheit	<ul style="list-style-type: none"> <li>• Aktualität der Informationen</li> <li>• Verständlichkeit der Informationen</li> <li>• Freitext- und Schlagwortsuche</li> <li>• Glossar für Sachbegriffe</li> <li>• Linksbündige Ausrichtung des Textes</li> <li>• Vermeidung von serifenhaltigen Schriften zur Verbesserung der Lesbarkeit</li> <li>• Vermeidung doppelten (vertikal + horizontal) Scrollens</li> <li>• Verwendung kurzer Sätze und zielgruppenspezifischer Wortwahl</li> <li>• Informationen über den Anbieter der Seite in ausreichender Tiefe</li> <li>• Konsistenz von Seitenlayout und Navigation</li> </ul>
Selbstbeschreibungsfähigkeit	<ul style="list-style-type: none"> <li>• Startseite sollte Kurzorientierung über Anbieter der Seite liefern</li> <li>• Existenz einer Sitemap zur Orientierung des Nutzers</li> <li>• Alternativtexte für Bilder und Plugins</li> </ul>
Kontrollierbarkeit	<ul style="list-style-type: none"> <li>• Aus- und einschalten von visuellen oder auditiven Medien</li> </ul>
Erwartungskonformität	<ul style="list-style-type: none"> <li>• Möglichkeit zur Kontaktaufnahme/Kontaktdatenangabe</li> <li>• „Sprechende Links“ in der Navigation</li> </ul>
Fehlertoleranz	<ul style="list-style-type: none"> <li>• Suchergebnisse werden auch bei fehlerhaften Eingaben in das Suchfeld geliefert</li> </ul>
Individualisierbarkeit	-
Lernförderlichkeit	<ul style="list-style-type: none"> <li>• Begrenzung der Textlänge (&lt;1,5 Seiten), wichtige Angaben sollten ohne Scrollen lesbar sein</li> </ul>
Ergänzender Kriterienkatalog	
	<ul style="list-style-type: none"> <li>• Vermeidung von Elementen mit langer Ladezeit oder hohen Systemanforderungen</li> <li>• Resize von Bildern, wenn möglich (Vermeidung langer Ladezeiten)</li> <li>• Angebot von Newsletter/RSS-Feeds</li> </ul>

## APPENDIX C – ITEMSÄTZE

### E-COMMERCE

<b>Aufgabenangemessenheit</b>
Der Warenkorb ist leicht auf der Seite zu finden
Das Vorhandensein eines Merkzettels erleichtert mir meinen Einkauf
Ich begrüße die Möglichkeit, Waren und Produkte nach ihren Eigenschaften (Preis, Relevanz) sortieren zu können
Die Anordnung der Produkte auf der Seite verwirrt mich
Die Informationen zu Lieferzeit und Versandkosten sind einfach zu finden
Die Hilfefunktion konnte meine Fragen klären
Wenn die Hilfefunktion nicht ausreichend genug ist, weiß ich, an wen ich mich wenden kann
Die Kontaktdaten zum Anbieter des Shops sind leicht auffindbar
Die Anzahl an Produktkategorien ist zu unübersichtlich
Ich weiß genau, unter welchem Navigationspunkt ich welche Produkte finde
Wenn ich ein bestimmtes Produkt suche kann ich es leicht über Suche und Navigation finden
Der Bestellvorgang dauert mir viel zu lang
<b>Selbstbeschreibungsfähigkeit</b>
Die Startseite wirkt übersichtlich und einladend
Ich weiß jederzeit, wo ich den Warenkorb finde
Ich weiß jederzeit, wo ich meinen Merkzettel finde
Die Suchfunktion ist leicht zu finden
Ich weiß immer, an welchem Punkt der Seite ich mich befinde
Ich weiß auf Anhieb, was meine Bestellung kosten wird
Ich weiß im Bestellvorgang, wie lange es voraussichtlich noch dauert.
Ich weiß manchmal nicht, ob ich eine Aktion erfolgreich durchgeführt habe
<b>Steuerbarkeit</b>
Produkte zum Warenkorb hinzuzufügen oder zu entfernen, ist leicht für mich
Über die Weitergabe meiner privaten Daten an Dritte mache ich mir Sorgen
Ich kann problemlos über den Browser vor- und zurücknavigieren
<b>Erwartungskonformität</b>
Die Navigation links bietet mir eine gute Orientierungshilfe



Die Schritte des Bestellvorgangs sind unkompliziert zu durchlaufen
Mein Kundenkonto ist übersichtlich dargestellt
Ich weiß, wie ich Angaben zu meinem Kundenkonto ändere
Ein- und Ausloggen ist für mich einfach
<b>Fehlertoleranz</b>
Die Anordnung der Produkte nach Ttigkeit einer Filtereinstellung (z.B. Sortierung nach Preis – aufsteigend) ist mir unerklrlich
Manchmal lsche ich Eingaben, weil ich versehentlich auf den Zurcksetzen-Button gekommen bin
Den „Eingaben zurcksetzen“-Button finde ich berflssig
<b>Individualisierbarkeit</b>
Die Empfehlungen des Systems zu anderen Produkten empfinde ich als lstig
Die individuellen Features nach dem Login (z.B. Begruung, Produktempfehlungen) finde ich sehr positiv
<b>Lernfrderlichkeit</b>
Die aktive Systemhilfe (z.B. Tour durch das System) empfinde ich als hilfreich
Ich finde mich auf Anhieb auf der Seite berhaupt nicht zurecht
<b>Ergnzend:</b>
Ich finde das Design der Seite ansprechend (z.B. Farbe, Layout)
Die Startseite verspricht mehr als die Seite wirklich hlt
Eine positive Verkuferbewertung strkt mein Vertrauen beim Kauf
Ich wrde die Produkte gerne verschlagworten (Tagging)
Die Verschlagwortung von Produkten stellt eine groe Hilfe fr mich dar
Ich habe mich durch eine Bewertung einer unabhngigen Institution oder Einrichtung (z.B. Stiftung Warentest) fr den Shop entschieden
Negative Bewertungen durch Verbrauchertests halten mich davon ab, in einem Shop zu kaufen
Ich empfinde Testberichte anderer Kunden als sehr hilfreich
Ich habe mich bereits auf Basis von Testberichten anderer Kunden fr ein Produkt entschieden
Die Autovervollstndigung meiner Eingaben in Formularfelder (z.B. Suchmaske) empfinde ich als hilfreich
Die Mglichkeit, Bilder von Produkten zu vergrern und/oder zu drehen, hilft mir bei der Kaufentscheidung
Abbildungen werden auf meinem Bildschirm manchmal nicht angezeigt

Ich kann den Kaufprozess im Online-Shop genauso gut durchführen wie in einem Geschäft

## E-LEARNING

<b>Aufgabenangemessenheit</b>
Ich finde mich auf der Benutzeroberfläche gut zurecht
Die verwendeten Wörter und Sätze sind für mich immer leicht zu verstehen
Es fällt mir leicht, mich auf den Lernstoff zu konzentrieren
Ich weiß, wie ich zwischen einzelnen Bereichen navigiere
Die Strukturierung der Lerninhalte ist für mich logisch
Die Verwendung von Abbildungen und Grafiken erleichtert mir den Lernprozess
Die Verwendung von Video- oder Audiodateien erleichtert mir den Lernprozess
<b>Selbstbeschreibungsfähigkeit</b>
Die Navigation ist übersichtlich gegliedert
Ich weiß, was sich hinter den einzelnen Navigationspunkten verbirgt
Die Navigation hilft mir, mich im Lernstoff zurechtzufinden
Wenn Symbole verwendet werden, ist mir klar, was sie bedeuten
Mir ist jederzeit klar, in welchem Modul ich mich befinde
Mir ist jederzeit klar, in welchem Schwierigkeitsgrad ich mich befinde
Manchmal ist mir unklar, wie ich von einem bestimmten Punkt aus an den Anfang zurückkomme
Mir ist klar, wo ich die Hilfestellung finde
Die Hilfestellung des Systems ist ausreichend für mich
<b>Steuerbarkeit</b>
Ich kann jederzeit problemlos zwischen den einzelnen Lerninhalten navigieren
Ein Wechsel der Schwierigkeitsstufe ist problemlos möglich
Der Wiedereinstieg in ein einmal abgebrochenes Modul ist unkompliziert
<b>Fehlertoleranz</b>
Fehlerhafte Eingaben im Nachhinein noch zu korrigieren, ist unkompliziert für mich
<b>Individualisierbarkeit</b>
Die Speicherung bereits absolvierter Lerninhalte ist eine große Hilfe für mich
Das System passt sich genau meinem Lernfortschritt an
Die individuelle Speicherung von Lernfortschritten finde ich überflüssig
Ich würde das Design gerne mehr meinen Bedürfnissen anpassen

Ich würde die Systemhilfe gerne ausschalten
<b>Lernförderlichkeit</b>
Die Inhalte sind meinem Lernfortschritt genau angepasst
Ich weiß auch nach längerer Zeit der Abwesenheit noch, wie man das System bedient
Die Hilfestellungen des Systems sind hilfreich
Ich wünsche mir eine Tour durch das System
Hinweise zur Lösung einer Aufgabe statt der Lösung selber finde ich gut
Texte würde ich mir lieber ausdrucken und auf Papier lesen

## ENTERTAINMENT

<b>Aufgabenangemessenheit</b>
Ich finde mich auf der Seite auf Anhieb zurecht
Die Funktionen der Seite sind mir auf Anhieb klar
Die Funktionen der Seite sind für mich einfach zu bedienen
Das Herunterladen von Plugins für die Seite empfinde ich als störend
Die Seite bietet auch nach längerer Benutzung noch Unterhaltungswert
Ich finde die Benutzung der Seite spannend und unterhaltsam
<b>Selbstbeschreibung</b>
Ich verstehe Icons und Symbole, die verwendet werden
<b>Steuerbarkeit</b>
Ich bin in der Lage, getätigte Einstellungen (Filter, Suchergebnisse) rückgängig zu machen
Mir ist klar, wo ich abspielende Mediendateien (Musik/Videos) unterbrechen und/oder anhalten kann
<b>Erwartungskonformität</b>
Die Gestaltung der Funktionalitäten (Abspielen von Mediendateien, Downloads) verwirrt mich
Das Design der Website gefällt mir
<b>Fehlertoleranz</b>
Wenn sich Dateien nicht abspielen lassen, ist mir klar, warum
<b>Individualisierbarkeit</b>
Persönliche Empfehlungen des Systems empfinde ich als störend
Die Größe von visuellen Elementen (Bildschirmen, Playern, Spielfenstern) lässt sich meinen Bedürfnissen ausreichend anpassen

<b>Lernförderlichkeit</b>
Auftretende Probleme lassen sich mithilfe der Systemhilfe lösen
Ich vermisste eine Tour durch das System
Durch die Guided Tour durch das Systems sind mir neue Funktionen klar geworden
Funktionen und Möglichkeiten des Systems sind mir ausreichend erklärt worden
Tutorials und Anleitungen sind so platziert, dass ich sie leicht finde
Auch nach mehrmaliger Benutzung verwirrt mich das System
<b>Eignung für das Kommunikationsziel</b>
Ich erkenne, wenn Elemente (Buttons, Schaltflächen) eine bestimmte Funktion haben
Manchmal suche ich vergeblich nach Funktionen
<b>Eignung für die Wahrnehmung</b>
Visuelle Elemente (Bildschirm, Icons, Thumbnails) sind für mich ausreichend groß
Ich empfinde die Anordnung des Seiteninhalts als übersichtlich
<b>Eignung für Exploration</b>
Die Gestaltung der Seite regt mich dazu an, auf ihr zu surfen (z.B. durch Empfehlungen)
Ich habe Spaß an der Interaktion mit der Seite
<b>Eignung für die Benutzungsmotivation</b>
Die Seite bietet auch nach längerer Nutzung noch neue Inhalte für mich
Mir wird schon nach kurzer Zeit auf der Seite langweilig
Die Seite wirkt originell auf mich
Mir gefällt die Vielzahl unterschiedlicher Aktivitäten, die ich auf der Seite durchführen kann
Die Funktionen der Seite fordern mich

## SOCIAL SOFTWARE

<b>Aufgabenangemessenheit</b>
Die Benutzeroberfläche ist übersichtlich und klar gehalten
Zur Kommunikation mit anderen Nutzern habe ich genug Funktionen
Die Kommunikationsfunktionen der Seite sind einfach zu bedienen

Ich habe Schwierigkeiten mit der Erstellung meines Nutzerprofils
<b>Selbstbeschreibungsfähigkeit</b>
Mir ist unklar, welche Funktionen die Seite eigentlich anbietet
Ich weiß, wo ich die allgemeinen Verhaltensregeln („Netiquette“) finde
Die Systemtour empfinde ich als hilfreich
Ich weiß bei Fragen jederzeit, wo die Hilfefunktion platziert ist
<b>Steuerbarkeit</b>
Ich weiß jederzeit, wo ich mein Profil ändern kann
Ich habe Probleme mit dem Erstellen, Editieren oder Löschen von Beiträgen
Die Verschlagwortung von Beiträgen (Tagging) ist nützlich für mich
Ich fühle mich durch eingeschränkte Sichtbarkeit meines Profils gegenüber Dritten sicherer
Ich empfinde eingeschränkte Sichtbarkeit anderer Nutzerprofile als störend
<b>Erwartungskonformität</b>
Die Gestaltung des Texteditors ist einfach zu verstehen
Vorgefertigte Emoticons (Smilies) empfinde ich als hilfreich
<b>Individualisierbarkeit</b>
Ich begrüße die Möglichkeit zur Individualisierung des Layouts (Farben, Design)
Ich würde mein Nutzerprofil gern ausführlicher beschreiben
Es ist mir möglich, nur so viele Informationen über mich preiszugeben, wie ich es möchte
<b>Lernförderlichkeit</b>
Eine Tour durch das System empfinde ich als hilfreich
Die Systemhilfe nehme ich oft in Anspruch
<b>Ergänzend</b>
Das Portal stellt eine wichtige Kommunikationsgrundlage für mich dar
Das Portal sollte 24 Stunden am Tag erreichbar sein
Die Möglichkeit, andere Nutzer zu bewerten, finde ich sinnvoll und nützlich
Die Existenz verschiedener Nutzergruppen (Anfänger, Fortgeschrittene) finde ich sinnvoll
Ich bin mir sicher, dass mit meinen privaten Daten verantwortlich umgegangen wird

## INFORMATION

<b>Aufgabenangemessenheit</b>
Ich fühle mich durch die Inhalte der Website immer auf dem neuesten Stand, was Informationen anbetrifft
Mir fehlen wichtige Themen, über die die Seite keine Information liefert
Die Informationen, die ich erhalte, sind glaubwürdig
Der Text auf der Seite ist gut lesbar (Schriftgröße, etc.)
Die Texte auf der Seite sind viel zu lang
Die Suchfunktion liefert mir die gewünschten Ergebnisse
Manche Wörter, die auf der Seite verwendet werden, verstehe ich nicht
Die Sätze, die in den Texten verwendet werden, sind mir viel zu lang
Die Benutzeroberfläche ist einfach und übersichtlich gehalten
<b>Selbstbeschreibungsfähigkeit</b>
Mir ist klar, welche Information ich wo finde
Ich weiß jederzeit, wie ich auf die Hauptseite komme
Mir fehlt eine Inhaltsangabe oder eine Zusammenfassung am Anfang des Textes
<b>Steuerbarkeit</b>
Ich weiß, wo ich multimediale Elemente (Audio/Videodateien) ein- oder ausschalte
Den Einsatz multimedialer Elemente empfinde ich als Bereicherung
<b>Erwartungskonformität</b>
Die Texte/Informationen sind für mich in erkennbarer Reihenfolge angeordnet
Die Texte sind inhaltlich gut strukturiert
<b>Individualisierbarkeit</b>
Für mich relevante Artikel werden mir empfohlen
Die Archivierung kürzlich angesehener Artikel ist für mich wichtig
<b>Lernförderlichkeit</b>
Mir fehlen Zusammenfassungen
Ich nutze die Newsletter- oder RSS-Funktion regelmäßig
Die Newsletter/RSS-Funktion empfinde ich als hilfreich
<b>Ergänzend</b>
Die Seite stellt für mich eine der wichtigsten Informationsquellen dar
Die Beiträge der Seite sind stets neutral formuliert

Die Seite verwendet zu viel Werbung
Die Möglichkeit, den Autor eines Textes zu kontaktieren, ist für mich wichtig
Ich drucke lange Texte meist aus, um sie zu lesen

## REPRÄSENTATION

<b>Aufgabenangemessenheit</b>
Über den Anbieter der Seite werden ausreichend Informationen geliefert
Die Informationen über den Anbieter der Seite sind stets auf dem neuesten Stand
Die Informationen, die über den Anbieter der Seite geliefert werden, sind für mich einfach zu verstehen
Die Texte auf der Seite sind einfach zu lesen (Schriftgröße etc.)
Mir ist klar, welche Informationen über den Anbieter der Seite ich in welchem Bereich finde
Die Seite verwendet zu viel Werbung
<b>Selbstbeschreibungsfähigkeit</b>
Die Startseite wirkt übersichtlich und ansprechend
Mir ist klar, was den Anbieter der Seite auszeichnet
Die Gliederung der Navigation empfinde ich als übersichtlich
<b>Steuerbarkeit</b>
Mir ist klar, wo ich Audio- oder Videodateien an- und abschalten kann
Den Einsatz audiovisueller Medien auf der Seite empfinde ich als hilfreich
<b>Erwartungskonformität</b>
Ich weiß, wie ich den Anbieter der Seite erreichen kann
<b>Lernförderlichkeit</b>
Ich fühle mich nach Besuch der Seite gut über den Anbieter informiert
Die Seite hatte keine neuen Informationen für mich
<b>Ergänzend</b>
Mein Computer benötigt lange, um Bilder und Seitenelemente zu laden
Der Aufbau der Seite wirkt unübersichtlich auf mich
Es werden dieselben Informationen mehrmals geliefert

### ANSCHREIBEN FÜR EXPERTEN

Hallo!

Vielen Dank, dass Sie sich Zeit nehmen, die folgenden Aussagen (Items) für einen Fragebogen für den Bereich E-Commerce zu beurteilen. Sie stellen einen Teil eines zu entwickelnden, adaptierbaren Tools dar, mit dessen Hilfe Fragebögen zur Beurteilung der Gebrauchstauglichkeit von Websites erstellt werden sollen. Es handelt sich hierbei um eine Abschlussarbeit im Projekt „Barrierefreiheit, Usability, Internationalisierung“, das die Universität Hildesheim in Zusammenarbeit mit der Firma usability.de betreibt.

Mit dem Bereich **E-Commerce** sind Websites gemeint, die online **Einkaufs- und Verkaufsmöglichkeiten** sowie Informationen über Produkte und Dienstleistungen für den Nutzer bereitstellen. Beispiele aus dem Bereich E-Commerce wären *amazon.de*, *ebay.de* oder *vv-computer.de*.

Die Items werden später im Fragebogen als 5-stufige Likert-Skala erscheinen

Beispiel:

	Stimme überhaupt nicht zu	Stimme weniger zu	Neutral	Stimme ziemlich zu	Stimme vollkommen zu
Der Warenkorb ist leicht auf der Seite zu finden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Bitte beurteilen Sie die einzelnen Items des Fragebogens vor allem unter dem Gesichtspunkt **Relevanz** (wie wichtig ist das abgefragte Merkmal in Bezug zum Anwendungskontext, hier E-Commerce?) sowie die komplette Skala auf das Kriterium **Vollständigkeit** (wurden alle wichtigen Merkmale abgefragt oder fehlt ein wichtiger Aspekt?).

Unter jedem Item ist ein Kommentarfeld für eventuelle Anmerkungen oder (Formulierungs-) Verbesserungsvorschläge. Sollten keine nötig sein, so können Sie das Feld leer lassen.

Die soziodemografischen Daten sind freiwillig und werden selbstverständlich vertraulich behandelt und nicht an Dritte weitergegeben.

Vielen Dank für Ihre Mithilfe!



## ANSCHREIBEN FÜR STUDIERENDE

### Aufgabe 1)

Beurteilen Sie die einzelnen Items eines möglichen Usability Fragebogens für Websites kritisch unter dem Gesichtspunkt **Verständlichkeit** (sind die einzelnen Aussagen klar und eindeutig formuliert oder gibt es Verständnisprobleme?). Versetzen Sie sich dabei bitte in die Rolle eines Nutzers ohne Vorwissen im Bereich Usability.

Bewerten Sie die einzelnen Items des Fragebogens weiterhin unter dem Gesichtspunkt **Relevanz** (wie wichtig ist das abgefragte Merkmal in Bezug zum Anwendungskontext, hier Online-Communities und Social Software?). Hierzu dienen die Checkboxes am Ende jeder Zeile.

Unter jedem Item ist ein **Kommentarfeld** für eventuelle Anmerkungen oder (Formulierungs-) Verbesserungsvorschläge. Sollten keine nötig sein, so können Sie das Feld leer lassen.

### Aufgabe 2)

Finden sie Aussagen, die nicht zur o.g. Website-Kategorie passen. Markieren Sie diese klar und deutlich, und begründen Sie ihre Auswahl!

### Hinweis:

Die Items werden später im Fragebogen als 5-stufige Likert-Skala erscheinen.

### Beispiel:

	Stimme überhaupt nicht zu	Stimme weniger zu	Neutral	Stimme ziemlich zu	Stimme vollkommen zu
Die Benutzeroberfläche ist übersichtlich und klar gehalten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die soziodemografischen Daten sind freiwillig und werden selbstverständlich vertraulich behandelt und nicht an Dritte weitergegeben.

## APPENDIX E: EXPERTEN UND STUDIERENDEN RELEVANZBEWERTUNG

s. Relevanz.xls auf der beiliegenden CD-ROM

## APPENDIX F: AUFLISTUNG DER ÜBERARBEITETEN ITEMS MIT RELEVANZBEWERTUNG

### E-COMMERCE

Warenkorb, Bestellung, Versand	Relevanzscore
Produkte zum Warenkorb hinzuzufügen ist leicht für mich	1
Es ist leicht, Produkte aus dem Warenkorb zu entfernen	1
Die Informationen zu Lieferzeit und Versandkosten sind einfach zu finden	0,97435897
Die Kosten der Bestellung (inkl. Versand) sind klar ersichtlich	0,94871795
Der Warenkorb ist leicht auf der Seite zu finden	0,93333333
Das System gibt ausreichend Rückmeldung bei durchgeführten Aktionen (z.B. Bestellung abschicken)	0,84102564
Der Bestellvorgang umfasst zu viele Schritte	0,78974359
Ich weiß immer genau, in welchem Schritt des Bestellvorgangs ich mich befinde	0,66666667
Ich kann problemlos im Kaufprozess vor und zurücknavigieren	0,60512821
Ich muss unnötig viele Angaben im Bestellvorgang machen	<sup>69</sup>
Ich weiß, welcher Klick den Kauf auslöst	

Produkte	Relevanzscore
Ich kann die Produkte meinen Wünschen entsprechend sortieren	0,97435897
Das Vorhandensein eines Merkzettels erleichtert mir meinen Einkauf	0,93333333
Die Möglichkeit, Bilder von Produkten zu vergrößern und/oder zu drehen, hilft mir bei der Kaufentscheidung	0,93333333
Ich empfinde Testberichte anderer Kunden als hilfreich	0,93333333

<sup>69</sup> Items ohne Relevanzangabe wurden neu hinzugefügt oder stellen überarbeitete oder umformulierte Items dar

Die Empfehlungen des Systems zu anderen Produkten sind für mich interessant	0,88205128
Ich weiß jederzeit, wo ich meinen Merkzettel finde	0,85641026
Die Kategorisierung der Produkte ist übersichtlich	0,81538462
Ich kann ein bestimmtes Produkt leicht über die Produktkategorien finden	0,81538462
Die Anordnung der Produkte auf der Seite ist übersichtlich	0,74871795
Abbildungen werden auf meinem Bildschirm manchmal nicht angezeigt	0,63076923

#### Suche & Hilfe

Die Suchfunktion ist leicht zu finden	0,97435897
Die Suchfunktion liefert mir nützliche Ergebnisse	0,93333333
Die Hilfefunktion ist ausreichend für mich	0,74871795
Die Hilfefunktion ist jederzeit auffindbar	0,71282051

#### Kundenkonto & Kundendaten

Mein Kundenkonto ist übersichtlich dargestellt	1
Ich weiß, wie ich Angaben zu meinem Kundenkonto ändere	0,93333333
Ein- und Ausloggen ist für mich einfach	0,83076923
Über die Weitergabe meiner privaten Daten an Dritte mache ich mir Sorgen	0,76410256

#### Seitenaufbau

Die Startseite wirkt einladend	0,93333333
Die Kontaktdaten zum Anbieter des Shops sind leicht auffindbar	0,88205128
Die Hauptnavigation stellt eine gute Orientierungshilfe dar	0,80512821

Die Seite wirkt auf Anhieb übersichtlich auf mich	0,76410256
---	------------

## ENTERTAINMENT

Funktionen	
Die Größe von visuellen Elementen (Playern, Spielfenstern) lässt sich meinen Bedürfnissen entsprechend anpassen	1
Mir ist klar, wo ich abspielende Mediendateien (Musik/Videos) unterbrechen und/oder anhalten kann	0,97777778
Die Funktionen der Seite sind einfach zu bedienen	0,95555556
Die Gestaltung von Funktionselementen (Eingabefelder, Buttons) ist selbsterklärend	0,75555556
Die Funktionen der Seite sind auf Anhieb ersichtlich	0,73333333
Auch nach mehrmaliger Benutzung verwirrt mich das System	0,71111111
Wenn sich Dateien (Musik/Videos) auf der Seite nicht abspielen lassen, ist mir klar, warum	0,68888889
Alle Funktionen sind leicht zu finden	0,64444444
Ich kann Einstellungen (Filter/Suchergebnisse) leicht rückgängig machen	0,46666667

Unterhaltungswert	
Die Gestaltung der Seite regt mich dazu an, auf ihr zu surfen (z.B. durch Empfehlungen)	0,95555556
Die Seite bietet auch nach längerer Benutzung noch Unterhaltungswert	0,91111111
Die Empfehlungen des Systems zu anderen Mediendateien (Musik/Videos) sind interessant	0,91111111
Die Seite wirkt originell auf mich	0,66666667
Mir gefällt die Vielzahl unterschiedlicher Aktivitäten, die ich auf der Seite durchführen kann	0,66666667
Ich habe Spaß an der Interaktion mit der Seite	0,51111111

Suche und Hilfe	
Die Hilfefunktion ist ausreichend für mich	0,82222222
Die Hilfefunktion ist jederzeit auffindbar	
Die Guided Tour konnte mir neue Funktionen zeigen oder erklären	0,57777778
Die Einführung/Anleitung zum Programm ist leicht zu finden	0,97777778

Design und Seitenaufbau	
Das Design der Website spricht mich an	0,95555556
Ich finde mich auf der Seite auf Anhieb zurecht	0,77777778
Ich verstehe Abbildungen und Symbole, die verwendet werden	0,77777778
Visuelle Elemente (Abbildungen, Grafiken) sind für mich ausreichend groß	0,75555556
Ich empfinde die Anordnung des Seiteninhalts als übersichtlich	0,73333333

## PORTALE

Funktionen & Kommunikation	
Die Funktionen zur Kommunikation mit anderen Nutzern sind leicht zu bedienen	1
Die Funktionen der Seite sind auf Anhieb ersichtlich	0,88095238
Das Portal stellt eine wichtige Kommunikationsgrundlage für mich dar	0,8452381
Die Möglichkeiten, mit anderen Nutzern zu kommunizieren, sind ausreichend	0,80952381
Das Erstellen, Editieren oder Löschen von Beiträgen ist problemlos möglich	0,79761905
Der Texteditor ist einfach zu bedienen	0,73809524
Die Möglichkeit zur Bewertung anderer Nutzer ist hilfreich	0,55952381
Die Kommentierung oder Verschlagwortung von Beiträgen ist einfach	

Nutzerprofil und Informationen	
Die Möglichkeit, das Nutzerprofil zu ändern, ist gut sichtbar platziert	1
Ich bin mir sicher, dass mit meinen privaten Daten verantwortlich umgegangen wird	0,97619048
Ich fühle mich durch eingeschränkte Sichtbarkeit meines Profils gegenüber Dritten sicherer	0,95238095
Es ist mir möglich, nur so viele Informationen über mich preiszugeben, wie ich es möchte	0,92857143
Das Nutzerprofil lässt sich leicht erstellen	0,86904762
Ich würde mein Nutzerprofil gern ausführlicher beschreiben	0,71428571
Die Existenz verschiedener Nutzergruppen (Anfänger, Fortgeschrittene) ist nützlich	0,51190476
Ein- und Ausloggen ist für mich einfach	

Suche und Hilfe	
Die Hilfefunktion ist ausreichend für mich	
Die Hilfefunktion ist jederzeit auffindbar	0,78571429
Die Suchfunktion ist leicht zu finden	
Die Suchfunktion liefert nützliche Ergebnisse	

Design und Seitenaufbau	
Die Benutzeroberfläche ist übersichtlich	0,75
Das Layout der Seite lässt sich meinen Bedürfnissen entsprechend anpassen	0,5952381
<i>Ich würde das Layout der Seite gerne meinen Bedürfnissen anpassen</i>	

## INFORMATION

Informationsgehalt	
Mir fehlen wichtige Themen, über die die Seite keine Information liefert	0,83333333
Die Informationen, die ich erhalte, sind glaubwürdig	0,97435897
Mir ist klar, welche Information ich wo finde	0,80769231
Inhaltsangaben und Zusammenfassungen sind nützlich	0,70512821
Durch die Informationen auf der Website fühle ich mich auf dem neuesten Stand	1
Die Texte/Informationen sind in sinnvoller Reihenfolge angeordnet	0,67948718
Artikel, die mir empfohlen werden, sind interessant für mich	0,94871795
<i>Ich wünsche mir Empfehlungen zu möglichen interessanten Artikeln</i>	
Die Seite stellt für mich eine wichtige Informationsquelle dar	0,44871795
Die Beiträge der Seite sind sachlich und neutral formuliert	0,44871795
Auf der Seite ist zu viel Werbung zu sehen	0,94871795
Die Seite wirkt seriös auf mich	

Textaufbau und -struktur	
Die Texte sind sinnvoll strukturiert	0,94871795
Die Texte sind angenehm auf dem Bildschirm zu lesen	0,94871795
Die Texte sind angemessen lang	0,67948718
Die Texte sind immer verständlich	0,32051282

Suche	
Die Suchfunktion ist leicht zu finden	
Die Suchfunktion liefert nützliche Ergebnisse	1

Benutzerfunktionen	
Die Newsletter/RSS-Funktion ist hilfreich	0,94871795
Ich drucke lange Texte meist aus, um sie zu lesen	0,76923077
Die Möglichkeit, den Autor eines Textes zu kontaktieren, ist für mich wichtig	0,58974359
Die Archivierung kürzlich angesehener Artikel ist für mich wichtig	0,3974359

Seitenaufbau und -elemente	
Die Benutzeroberfläche ist einfach und übersichtlich gehalten	0,80769231
Den Einsatz multimedialer Elemente (Audio/Videodateien) empfinde ich als Bereicherung	0,80769231
Ich weiß, wo ich multimediale Elemente (Audio/Videodateien) ein- oder ausschalte	0,73076923

## REPRÄSENTATION

Informationsgehalt	
Die Texte sind angenehm auf dem Bildschirm zu lesen	1
Ich fühle mich nach Besuch der Seite gut über den Anbieter informiert	0,97777778
Die Informationen über den Anbieter der Seite sind auf dem neuesten Stand	0,93333333
Über den Anbieter der Seite werden ausreichend Informationen geliefert	0,86031746
Die Informationen über den Anbieter der Seite sind leicht verständlich	0,83809524
Auf der Seite ist zu viel Werbung zu sehen	0,55873016
Dieselben Informationen werden unnötig oft geliefert	0,43492063
Die Seite wirkt seriös auf mich	
Die Texte sind angemessen lang	



Seitenaufbau und -elemente	
Die Kontaktdaten zum Anbieter der Seite sind leicht zu finden	1
Die Hauptnavigation bietet eine gute Orientierungshilfe	0,88253968
Den Einsatz multimedialer Elemente (Audio/Videodateien) auf der Seite empfinde ich als Bereicherung	0,77142857
Mein Computer benötigt lange, um Bilder und Seitenelemente zu laden	0,7047619
Ich weiß, wo ich multimediale Elemente (Audio/Videodateien) der Seite ein- oder ausschalte	0,67619048

Suche	
Die Suchfunktion ist leicht zu finden	
Die Suchfunktion liefert nützliche Ergebnisse	

## E-LEARNING

Lerninhalte	
Die Lerneinheiten sind sinnvoll strukturiert	0,97619048
Die Korrektur fehlerhafter Eingaben ist im Nachhinein problemlos möglich	0,97619048
Die Speicherung bereits absolvierter Lerninhalte ist hilfreich	0,97619048
Der Wiedereinstieg in eine zuvor beendete oder abgebrochene Lerneinheit ist einfach	0,95238095
Texte würde ich mir lieber ausdrucken und auf Papier lesen	0,92857143
Ich kann die Schwierigkeitsgrade meinen Bedürfnissen entsprechend anpassen	0,86666667
Ein Wechsel der Schwierigkeitsstufe ist problemlos möglich	0,81904762
Wenn Symbole verwendet werden, ist mir klar, was sie bedeuten	0,81904762
Ich weiß auch nach längerer Zeit der Abwesenheit noch, wie man das System bedient	0,81904762

Zusätzliche Hinweise statt der sofortigen Lösung der Aufgabe sind hilfreich beim Lernen	0,70952381
Die Aufmachung der Seite lenkt vom Lernstoff ab	0,66190476
Die Texte in den Lerneinheiten sind angemessen lang	
Die Texte der Lerneinheiten sind verständlich	
<i>Absolvierte Lerninhalte würde ich gerne speichern</i>	
<i>Ich wünsche mir zusätzliche Hinweise statt der sofortigen Lösung der Aufgabe</i>	

Navigation	
Mir ist jederzeit klar, in welchem Schwierigkeitsgrad ich mich befinde	0,97619048
Mir ist jederzeit klar, in welcher Lerneinheit ich mich befinde	0,92857143
Manchmal ist mir unklar, wie ich an den Anfangspunkt einer Lerneinheit zurückkomme	0,9047619
Ich kann problemlos zwischen Lerneinheiten navigieren	0,86666667
Die Hauptnavigation bietet eine gute Orientierungshilfe im Lernstoff	0,84285714
Ich weiß, welche Lerneinheiten wo zu finden sind	0,70952381

Seitenaufbau und -elemente	
Die Benutzeroberfläche ist übersichtlich	0,86666667
Die verwendeten Abbildungen und Grafiken erleichtern mir das Lernen	0,84285714
Die verwendeten Video- oder Audiodateien erleichtern mir das Lernen	0,70952381
Ich kann das Layout der Seite meinen Bedürfnissen entsprechend anpassen	0,45714286
<i>Ich würde das Design gerne meinen Bedürfnissen anpassen</i>	
Die Texte der Lerneinheiten sind angenehm auf dem Bildschirm zu lesen	

Hilfe	
Die Hilfefunktion zu den Lerneinheiten ist ausreichend für mich	0,97619048
Die Hilfefunktion zu den Lerneinheiten ist jederzeit auffindbar	0,70952381
Ich wünsche mir eine Tour durch das System	0,70952381
<i>Die Tour durch das System ist hilfreich</i>	

## APPENDIX G – HISTOGRAMME UND TABELLEN DER SPSS-AUSWERTUNG

s. Auswertung.xls auf der beiliegenden CD-ROM

## APPENDIX H – PROTOTYP

s. Prototyp.rar auf der beiliegenden CD-ROM

## APPENDIX I – HALBIERUNGSRELIABILITÄT DER KATEGORIE ENTERTAINMENT

[DatenSet1] F:\Downloader\SPSSII\Entertainment.sav

### Zusammenfassung der Fallverarbeitung

	N	%
Fälle Gültig	43	100,0
Ausgeschlossen <sup>a</sup>	0	,0
Gesamt	43	100,0

a. Listenweise Löschung auf der Grundlage aller Variablen in der Prozedur.

### Reliabilitätsstatistiken

Cronbachs Alpha	Teil 1	Wert	1,000
		Anzahl der Items	1 <sup>a</sup>
	Teil 2	Wert	1,000
		Anzahl der Items	1 <sup>b</sup>
		Gesamtzahl der Items	2
Korrelation zwischen Formen			,118
Spearman-Brown-Koeffizient	Gleiche Länge		,211
	ungleiche Länge		,211
Guttmans Split-Half-Koeffizient			,032

a. Die Items sind: Id

b. Die Items sind: V3

## EIGENSTÄNDIGKEITSERKLÄRUNG NACH §26 ABS. 6 DER MAGISTERPRÜFUNGSORDNUNG

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig abgefasst und nicht anderweitig zu Prüfungszwecken verwendet habe. Weiterhin erkläre ich, dass ich die Arbeit ausschließlich unter Verwendung der angegebenen Quellen und Hilfsmittel erstellt und alle wörtlichen und sinngemäßen Zitate aus diesen Quellen geeignet gekennzeichnet habe.

Hildesheim, den 01. April 2009

Franziska Leithold